

UNIVERSIDADE FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

HENRIQUE PRETI

**CARACTERIZAÇÃO FUNCIONAL DE GENES
CODIFICADORES DE PROTEÍNAS HIPOTÉTICAS
DIFERENCIALMENTE EXPRESSOS DURANTE A
METACICLOGÊNESE DE *TRYPANOSOMA CRUZI***

CURITIBA

2007

HENRIQUE PRETI

**CARACTERIZAÇÃO FUNCIONAL DE GENES
CODIFICADORES DE PROTEÍNAS HIPOTÉTICAS
DIFERENCIALMENTE EXPRESSOS DURANTE A
METACICLOGÊNESE DE *TRYPANOSOMA CRUZI***

Dissertação apresentada ao departamento de
Biologia Celular e Molecular da Universidade
Federal do Paraná, como pré-requisito na obtenção
do título de Mestre em Biologia Celular e Molecular.

Orientador: Dr. Marco Aurélio Krieger

Co-orientador: Dr. Christian Macagnan Probst

CURITIBA

2007


PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Departamento de Biologia Celular e Departamento de Fisiologia
Setor de Ciências Biológicas
Universidade Federal do Paraná
Instituto de Biologia Molecular do Paraná

DECLARAÇÃO

Declaramos para os devidos fins que Henrique Preti, no dia 13 de agosto de 2007, defendeu sua Dissertação de Mestrado em Biologia Celular e Molecular intitulada: "*Caracterização funcional de genes diferencialmente expressos durante a metaciclogênese de Trypanosoma cruzi codificadores de proteínas hipotéticas*", com a Banca Examinadora constituída pelos professores: Dr. Marco Aurélio Krieger, orientador do Instituto de Biologia Molecular do Paraná - IBMP, Dr. Marcelo Távora Mira, da Pontifícia Universidade Católica do Paraná - PUCPR e Dr. Milton Ozório Moraes, da Fundação Oswaldo Cruz - FIOCRUZ. Tendo sido Aprovado recebendo o título de Mestre em Biologia Celular e Molecular, área de Concentração em Biologia Celular e Molecular, desde que apresente a versão definitiva da dissertação conforme regimento interno do Programa.

Curitiba, 13 de agosto de 2007.


Prof. Dr. Ricardo Fernandez Perez
Coordenador do Programa de Pós-Graduação
Biologia Celular e Molecular

Aos meus pais e minha amada,

a quem dedico todos meus passos.

Agradecimentos

Agradeço a todos meus amigos, colegas e profissionais que me acompanharam ao longo desta curta jornada, preparando-me para a próxima.

Agradeço a diretoria do Instituto de Biologia Molecular do Paraná (IBMP), por depositar confiança em mim e me proporcionar tal aprendizado.

A todos amigos os quais foram surgindo ao longo da minha vivência no IBMP.

Agradeço em especial ao Dr. Marco Aurélio Krieger, Dr. Christian Macagnan Probst e Msc. Viviane Monteiro Góes, que conduziram com sapiência meus primeiros passos na ciência.

Sumário

Agradecimentos	iv
Sumário	v
Lista de Figuras	viii
Lista de Tabelas	xi
Resumo	xii
Abstract	xiii
1 Introdução	1
1.1 O Trypanosoma cruzi e o ciclo evolutivo.....	1
1.2 A doença de chagas	4
1.3 Características genômicas de T. cruzi	5
1.4 O Trypanosoma cruzi e a transcrição policistrônica	8
1.5 A regulação da expressão gênica em T. cruzi	11
1.6 A metaciclogênese.....	13
1.7 O estudo genômico funcional de T. cruzi.....	16
1.8 Proteínas hipotéticas em T. cruzi.....	19
2 Objetivos.....	21
2.1 Objetivo geral	21
2.2 Objetivos Específicos	21
3 Justificativas	22
4 Material e Métodos.....	24
4.1 Confecção dos microarranjos de DNA.....	24
4.2 Arquitetura do microarranjo	24
4.3 Desenho experimental	25
4.4 Fluxograma das etapas desenvolvidas.....	27
4.5 Análise dos dados gerados pelo microarranjo de DNA e seleção dos genes ...	27
4.6 Amplificação dos genes.....	36
4.7 Purificação dos produtos de PCR.....	36
4.8 Clonagem e vetores	37
4.8.1 Obtenção dos clones de entrada.....	39
4.8.2 Obtenção dos clones de expressão.....	40

4.9	Expressão das proteínas.....	41
4.10	Comprovação da expressão	42
4.11	Purificação das proteínas	42
4.12	Obtenção dos anticorpos policlonais.	44
4.13	Verificações da qualidade dos soros obtidos utilizando ensaios de Western blot.	44
4.14	Ensaio de Western blot para determinação de expressão	45
4.15	Análise dos resultados obtidos a partir dos ensaios de Western blot.	46
4.16	Imunolocalização por microscopia óptica	46
5	Resultados.....	48
5.1	Descrição inicial	48
5.2	Genes selecionados a partir das análises de microarranjo de DNA	48
5.3	Amplificação dos genes e purificação dos produtos	49
5.4	Obtenção dos clones de entrada.....	50
5.5	Obtenção dos clones de expressão.....	51
5.6	Sequenciamento dos clones obtidos	52
5.7	Expressão e purificação das proteínas	53
5.8	Obtenção dos anticorpos policlonais	55
5.9	Imunolocalização por microscopia óptica	55
5.10	Características gerais dos genes selecionados e resultados de western blot e imunolocalização.....	56
5.10.1	TcExp02A01, 8530.t00003, Proteína hipotética conservada	56
5.10.2	TcExp02A08, 8694.t00004, Proteína Hipotética	61
5.10.3	TcExp02A09, 7626.t00009, Proteína hipotética conservada	62
5.10.4	TcExp02A10, 6680.t00001, Sterol 24-c-methyltransferase	65
5.10.5	TcExp02A02, 6996.t00063, Proteína Hipotética Conservada.....	68
5.10.6	TcExp02A11, 8773.t00014, Proteína Hipotética	71
5.10.7	TcExp02A12, 8364.t00002, Proteína Hipotética conservada	72
5.10.8	TcExp02B01, 8208.t00006, Ras-related protein rab-5.....	73
5.10.9	TcExp02A03, 7109.t00005, Proteína hipotética conservada	76
5.10.10	TcExp02B02, 8764.t00016, Proteína Hipotética.....	77
5.10.11	TcExp02B03, 8728.t00021, Proteína Hipotética Conservada.....	78
5.10.12	TcExp02B04, 8257.t00019, DREV Metiltransferase.....	80

5.10.13	TcExp02A04, 6896.t00027, Proteína Hipotética conservada	81
5.10.14	TcExp02B05, 7224.t00006, Proteína Hipotética.....	85
5.10.15	TcExp02B06, 6958.t00014, Proteína Hipotética Conservada.....	85
5.10.16	TcExp02B07 8243.t00004, Dynein light chain 2B, cytoplasmic, putative	87
5.10.17	TcExp02A05, 4715.t00001, Anti-silencing protein, ASF1-like.....	89
5.10.18	TcExp02B08, 5595.t00003, Proteína hipotética.....	91
5.10.19	TcExp02B09, 8171.t00006, Proteína Hipotética conservada.....	92
5.10.20	TcExp02B10, 8416.t00006, Calcineurin B subunit, putative.....	94
5.10.21	TcExp02A06, 8415.t00003, Proteína hipotética conservada	97
5.10.22	TcExp02B11, 7000.t00002, Proteína hipotética.....	99
5.10.23	TcExp02B12, 8318.t00003, Proteína Hipotética conservada GTPase ...	100
5.10.24	TcExp02C01, 7168.t00007, Syntaxin putative.....	103
5.10.25	TcExp02A07, 5323.t00004, Proteína hipotética conservada	106
5.10.26	TcExp02C02, 8359.t00034, Proteína Hipotética Conservada.....	108
5.10.27	TcExp02C03, 7741.t00008, Proteína Hipotética Conservada.....	110
5.10.28	TcExp02C04, 6996.t00038, Centrin Putative.....	110
5.11	Consolidação dos resultados laboratoriais.....	112
5.12	Análise comparada do transcriptoma e proteômica de T. cruzi	115
6	Discussão.....	118
7	Conclusões.....	138
8	Perspectivas	140
9	Referências bibliográficas	142

Lista de Figuras

Figura 1.1 Ciclo de vida de <i>T. cruzi</i> . (modificado de www.dpd.cdc.gov/dpdx).	2
Figura 4.1 Desenho experimental das hibridações realizadas.	26
Figura 4.2. Fluxograma descritivo das etapas laboratoriais realizadas.	27
Figura 4.3 Agrupamento hierárquico dos dados de microarranjo disponíveis no IBMP.	30
Figura 4.4 Cluster contendo a sonda referente ao gene 4715.t00001, grupo 5.	31
Figura 4.5 Cluster contendo a sonda referente ao gene 5323.t00004, grupo 7.	31
Figura 4.6 Cluster contendo a sonda referente ao gene 6896.t00027, grupo 4.	32
Figura 4.7 Cluster contendo a sonda referente ao gene 6996.t00063, grupo 2.	32
Figura 4.8 Cluster contendo a sonda referente ao gene 7109.t00005, grupo 3.	33
Figura 4.9 Cluster contendo a sonda referente ao gene 8415.t00003, grupo 6.	33
Figura 4.10 Cluster contendo a sonda referente ao gene 8530.t00003, grupo 1.	34
Figura 4.11 Possibilidades de utilização do sistema Gateway a partir da obtenção do clone de entrada.	37
Figura 4.12 Reação attB x attP.	38
Figura 4.13 Reação attL x attR.	38
Figura 4.14 Vetor de entrada utilizado, pDONR 221(Invitrogen).	39
Figura 4.15 Vetor de expressão pDEST TM 17 (Invitrogen).	39
Figura 5.1 Padrão eletroforético obtido a partir das amostras de pcr purificadas.	50
Figura 5.2 Perfil eletroforético dos produtos obtidos da PCR de colônia de clones de entrada.	51
Figura 5.3 Perfil eletroforético dos clones de entrada purificados.	51
Figura 5.4 Perfil eletroforético dos produtos obtidos da PCR de colônia de clones de destinação.	52
Figura 5.5 Perfil eletroforético dos clones de expressão purificados.	52
Figura 5.6 Perfil eletroforético (SDS-PAGE) contendo o gene ID.	54
Figura 5.7 Gráfico LogoPlot representando o padrão de conservação do domínio quinase.	57
Figura 5.8. Comparação entre TcExp02A01 e o domínio ALBA.	58
Figura 5.9 Análise de similaridade entre TcExp02A01 e ortólogos de outros eucariotos.	58
Figura 5.10 Análise de similaridade entre TcExp02A01 e ortólogos de protozoários.	59
Figura 5.11 Árvore filogenética das proteínas contendo os domínios ALBA e RGG-box.	60
Figura 5.12 Análise de expressão da proteína TcExp02A01 por western blot em extratos de <i>T. cruzi</i> .	60
Figura 5.13 Imunolocalização de TcExp02A01.	61
Figura 5.14 Comparação entre TcExp02A08 e 7963.t00006.	62
Figura 5.15 Comparação entre TcExp02A09 e o domínio sacaropina desidrogenase.	63
Figura 5.16 Comparação entre a proteína codificada por TcExp02A09 e os ortólogos encontrados.	63
Figura 5.17 Análise de similaridade entre TcExp02A09 e os ortólogos em <i>G. violaceus</i> e <i>P. troglodytes</i> .	64
Figura 5.18. Análise de expressão da proteína TcExp02A09 por western blot em extratos de <i>T. cruzi</i> .	64
Figura 5.19 Imunolocalização da proteína TcExp02A09.	65
Figura 5.20 Comparação entre TcExp02A10 e o domínio Sterol_MT_C e Methyltransf_11.	66

Figura 5.21 Análise de similaridade entre seqüências de <i>T. cruzi</i> e outros eucariotos.	67
Figura 5.22. Análise de expressão da proteína TcExp02A10 por western blot em extratos de <i>T. cruzi</i> .	67
Figura 5.23 Imunolocalização da proteína TcExp02A10.	68
Figura 5.24 Gráfico LogoPlot representando o padrão de conservação do domínio quinase.	69
Figura 5.25. Comparação entre a seqüência de <i>T. cruzi</i> e o domínio PKinase.	70
Figura 5.26 Análise de similaridade entre seqüências de <i>T. cruzi</i> e outros eucariotos.	71
Figura 5.27 Análise de similaridade entre seqüências de <i>T. cruzi</i> e outros tripanossomatídeos.	73
Figura 5.28. Comparação entre TcExp02B01 e o domínio RAS	74
Figura 5.29 Análise de similaridade entre TcExp02B01 e ortólogos de outros eucariotos.	74
Figura 5.30. Análise de expressão da proteína TcExp02B01 por western blot em extratos de <i>T. cruzi</i> .	75
Figura 5.31 Imunolocalização da proteína TcExp02B01.	76
Figura 5.32. Comparação entre TcExp02A03 e o domínioFKBP_C.	76
Figura 5.33 Análise de similaridade entre TcExp02A03 e ortólogos de outros tripanossomatídeos.	77
Figura 5.34 Análise de similaridade entre seqüências de <i>T. cruzi</i> e <i>L. infantum</i> .	78
Figura 5.35 Gráfico HMM relativo o domínio zf-C2H2.	79
Figura 5.36 Comparação entre a seqüência de <i>T. cruzi</i> e o domínio zf-C2H2.	79
Figura 5.37 Análise de similaridade entre seqüências de <i>T. cruzi</i> e outros eucariotos.	80
Figura 5.38 Comparação entre a seqüência de <i>T. cruzi</i> e o domínio DREV.	81
Figura 5.39 Análise de similaridade entre TcExp02B04 e os ortólogos de outros eucariotos.	81
Figura 5.40 Comparação entre a seqüência de <i>T. cruzi</i> e os domínios SWIM e zf-C3HC4 utilizando o programa PFAM.	82
Figura 5.41. Gráfico HMM relativo ao domínio SWIM	83
Figura 5.42. Gráfico HMM do domínio zf-C3HC4 (RING).	83
Figura 5.43 Análise de similaridade entre TcExp02A04 e ortólogos de outros eucariotos.	84
Figura 5.44. Análise de expressão da proteína TcExp02A04 por western blot em extratos de <i>T. cruzi</i> .	84
Figura 5.45. Alinhamento das proteínas de TcExp02B05 presentes em CL Brener.	85
Figura 5.46 Gráfico HMM do domínio zf-CSL.	86
Figura 5.47 Comparação entre a seqüência de <i>T. cruzi</i> e o domínio zf-CSL.	86
Figura 5.48 Análise de similaridade entre seqüências de <i>T. cruzi</i> e outros eucariotos.	87
Figura 5.49 Análise de expressão da proteína TcExp02B06 por western blot em extratos de <i>T. cruzi</i> .	87
Figura 5.50. Comparação entre a seqüência de <i>T. cruzi</i> e o domínio Roadblock/LC7.	88
Figura 5.51. Análise de similaridade entreTcExp02B07 e ortólogos de outros eucariotos.	88
Figura 5.52. Análise de expressão da proteína TcExp02B06 por western blot em extratos de <i>T. cruzi</i> .	88
Figura 5.53 Imunolocalização de TcExp02B07.	89
Figura 5.54. Comparação entre a seqüência de <i>T. cruzi</i> e os domínios Anti-silence.	89
Figura 5.55. Gráfico HMM do domínio Anti-silence.	90
Figura 5.56. Análise de similaridade entre TcExp02A05 e ortólogos de outros eucariotos.	90
Figura 5.57. Análise de similaridade entre duas seqüências de <i>T. cruzi</i> CL Brener.	91
Figura 5.58. Análise de similaridade entre seqüências de <i>T. cruzi</i> , <i>T. congolense</i> e <i>T. vivax</i> .	92
Figura 5.59. Gráfico HMM do domínio PDC2-C.	93

Figura 5.60. Comparação entre a sequência de <i>T. cruzi</i> e o domínio PDC2_C.	93
Figura 5.61. Análise de similaridade entre TcExp02B09 e ortólogos de outros eucariotos.	94
Figura 5.62. Gráfico HMM do domínio EF-hand.	95
Figura 5.63. Comparação entre a sequência de <i>T. cruzi</i> e o domínio EF-hand	95
Figura 5.64. Análise de similaridade entre TcExp02B10 e ortólogos de outros eucariotos.	96
Figura 5.65. Análise de expressão da proteína TcExp02B10 por western blot em extratos de <i>T. cruzi</i> .	96
Figura 5.66. Imunolocalização de TcExp02B10.	97
Figura 5.67. Gráfico HMM retirado do programa PFAM mostrando o grau de conservação em cada posição aminoacídica do domínio PPR.	97
Figura 5.68. Comparação entre a sequência de <i>T. cruzi</i> e o domínio PPR.	98
Figura 5.69. Análise de similaridade entre sequências de <i>T. cruzi</i> e outros eucariotos.	98
Figura 5.70. Análise de similaridade entre sequências de <i>T. cruzi</i> e outros tripanossomatídeos.	99
Figura 5.71. Árvore filogenética mostrando as relações evolutivas da família de TcExp02B11.	100
Figura 5.72. Gráfico HMM do domínio ArfGap.	101
Figura 5.73. Comparação entre a sequência de <i>T. cruzi</i> e o domínio ArfGap.	101
Figura 5.74. Análise de similaridade entre TcExp02B12 e ortólogos de outros eucariotos.	102
Figura 5.75. Análise de expressão da proteína TcExp02B12 por western blot em extratos de <i>T. cruzi</i> .	102
Figura 5.76. Imunolocalização de TcExp02B12	103
Figura 5.77. Gráfico HMM do domínio SNARE.	104
Figura 5.78. Comparação entre a sequência de <i>T. cruzi</i> e o domínio SNARE.	104
Figura 5.79. Análise de similaridade entre sequências de <i>T. cruzi</i> e outros eucariotos.	105
Figura 5.80. Análise de expressão da proteína TcExp02C01 por western blot em extratos de <i>T. cruzi</i> .	105
Figura 5.81. Imunolocalização da proteína relativa ao gene 7168.t00007.	106
Figura 5.82. Comparação entre TcExp02A07 e o domínio PKinase.	107
Figura 5.83. Análise de similaridade TcExp02A07 e ortólogos de outros eucariotos.	107
Figura 5.84. Gráfico HMM do domínio TLD.	108
Figura 5.85. Comparação entre a TcExp02C02 e o domínio TLD.	109
Figura 5.86. Análise de similaridade entre TcExp02C02 e ortólogos de outros tripanossomatídeos.	109
Figura 5.87. Análise de similaridade entre sequências de <i>T. cruzi</i> e <i>T. cruzi</i> .	110
Figura 5.88. Comparação entre a sequência de <i>T. cruzi</i> e o domínio EF-hand.	111
Figura 5.89. Análise de similaridade entre TcExp02C04 e ortólogos de outros eucariotos.	111
Figura 5.90. Análise de expressão da proteína TcExp02C04 por western blot em extratos de <i>T. cruzi</i> .	112
Figura 5.91. Imunolocalização de TcExp02C04.	112
Figura 5.92. Gráfico heatmap mostrando a correlação entre transcriptoma e proteoma.	116
Figura 6.1. Imunolocalização de TcDhh1 e TcExp02A01.	133

Lista de Tabelas

<i>Tabela 4.1. Descrição dos genes selecionados para caracterização e sua classificação interna</i>	<i>29</i>
<i>Tabela 4.2: Iniciadores construídos adaptados para plataforma Gateway[®].</i>	<i>34</i>
<i>Tabela 5.1: Genes selecionados para caracterização.</i>	<i>49</i>
<i>Tabela 5.2 Resultados da expressão de proteínas na fase de teste e produção.</i>	<i>54</i>
<i>Tabela 5.3 Tabela dos genes dos quais obteve-se anticorpo policlonal.</i>	<i>55</i>
<i>Tabela 5.4. Eficiência de cada etapa de análise para as 28 proteínas analisadas.</i>	<i>113</i>

Resumo

A análise por microarranjo da metaciclogênese de *Trypanosoma cruzi* resultou na identificação de um conjunto de aproximadamente 1.000 genes diferencialmente expressos, em sua maioria proteínas hipotéticas. Visando aumentar o conhecimento sobre a função desses genes como a compreensão da expressão, função e regulação dos mesmos, com ênfase nas proteínas hipotéticas, iniciou-se o presente trabalho. Inicialmente, sete proteínas hipotéticas conservadas foram selecionadas, com base em sua confiabilidade de expressão diferencial. Posteriormente, foi realizada uma clusterização hierárquica com diferentes dados de metaciclogênese, para aumentar o grau de informação com relação aos padrões de expressão. Dentro de cada cluster dos genes previamente selecionados, foram identificados mais três genes distintos, codificando outra proteína hipotética conservada (gene secundário), uma proteína hipotética (gene acessório) e uma proteína com função conhecida (gene guia). Esses critérios visam aumentar a quantidade de proteínas hipotéticas sendo caracterizadas, bem como ter elementos de referência para futuras análises. Estes 28 genes foram clonados em um vetor de entrada (pDONR; Invitrogen), a partir do qual é possível inserir o gene de interesse em outros vetores, os quais apresentam propriedades diferentes. Inicialmente, realizamos a expressão em *E. coli* utilizando vetores de destinação (pDEST17, Invitrogen) para obtenção de proteínas recombinantes. Dos 28 genes selecionados foram obtidos 27 clones de entrada, os quais foram totalmente transferidos para o pDEST17. Na etapa de expressão, foram obtidas 19 proteínas, todas insolúveis, procedendo-se à obtenção de anticorpos policlonais, através da inoculação em camundongos, sendo que obteve-se soros para o reconhecimento de 18 proteínas recombinantes. Ao se realizar *western blot* para a comprovação da expressão diferencial, a proteína nativa foi identificada adequadamente por onze anticorpos, e o padrão de expressão diferencial, juntamente com outros 12 soros existente no instituto, em sua grande maioria corrobora os resultados do microarranjo. Foram verificados padrões de imunolocalização para 9 proteínas sendo 3 hipotéticas conservadas e 6 de função conhecida. Outros resultados obtidos por este trabalho foi o incremento nas informações sobre os genes, com relação a aspectos bioinformáticos, possibilitando uma melhor anotação e compreensão dos mesmos.

Abstract

Microarray analysis of *T. cruzi* metacyclogenesis resulted in the identification of approximately 1,000 differentially expressed genes and several of them encode hypothetical proteins. Aiming to increase the knowledge about the function of these genes, a better understanding of the expression, function and regulation, with an emphasis on hypothetical proteins, we have devised this work. Initially, seven conserved hypothetical proteins were selected on basis of differential expression confiability. Subsequently, hierarchical clustering was applied to the metacyclogenesis data and, within each cluster of genes previously selected, were identified other three distinct genes, encoding another conserved hypothetical protein (secondary gene), a hypothetical protein (accessory gene) and a protein with known function (guide gene). This approach increases the expectation of having at least one hypothetical protein better characterized, and also provides reference elements for future analysis. These 28 genes were cloned in entry vector of the Gateway platform (pDONR; Invitrogen), from which it is possible to insert the selected gene in other vectors, with different functional characterization capabilities. Initially, the 28 selected genes were expressed as proteins in *E. Coli* using appropriate destination vectors (pDEST17, Invitrogen), obtaining a recombinant protein. Of the 28 selected genes, 27 entry clones were obtained, which have been transferred to pDEST17. In the expression stage, we have obtained 19 proteins, all insoluble, which were used for polyclonal antibody production through inoculation in mice, and we have obtained sera recognizing 18 recombinant proteins. Western blot assay was performed aiming the differential expression analysis. The native protein was identified properly by 11 antibodies, and the differential expression pattern, including other 12 sera available in the Institute, mostly corroborates the microarray results. Nine protein were immunolocalized, 3 hypothetical protein and 6 of known function. Other results of this work were a better annotation, using bioinformatics tools, producing an increase in the functional knowledge of these proteins.

1 Introdução

1.1 O *Trypanosoma cruzi* e o ciclo evolutivo

O *Trypanosoma cruzi* (Chagas, 1909) é um protozoário parasita pertencente ao reino Protista, sub-reino Protozoa, filo Sarcomastigophora, sub-filo Mastigophora, classe Zoomastigophora, ordem Kinetoplastida, família Trypanosomatidae, gênero *Trypanosoma* e espécie *Trypanosoma cruzi* (LEVINE *et al*, 1980).

Os organismos da ordem Kinetoplastida apresentam uma especialização denominada cinetoplasto, correspondendo a sua mitocôndria única, e composto por uma grande quantidade de DNA em forma de bastão, o qual é denominado kDNA. Um dos critérios para se definir as formas evolutivas presentes na família Trypanosomatidae é a posição do cinetoplasto em relação ao núcleo (HOARE & WALLACE, 1966; MASLOW & SIMPSON, 1995). No *T. cruzi*, as formas epimastigotas possuem o cinetoplasto em forma de bastão e em posição anterior ao núcleo, constituído por um material filamentosso organizado em feixes de fibras fortemente empacotados e perpendicularmente orientados ao eixo longitudinal do parasita. Já as formas tripomastigotas possuem o cinetoplasto na região posterior, com uma forma mais arredondada devido à diminuição no grau de empacotamento do kDNA. As formas amastigotas apresentam o cinetoplasto de maneira muito semelhante aos epimastigotas (SOARES, 1989).

O ciclo evolutivo do *T. cruzi* consiste na alternância entre dois vetores, um inseto hematófago da família Reduviidae e um hospedeiro mamífero. Didaticamente, o ciclo é iniciado quando o hospedeiro vertebrado é infectado pela forma tripomastigota metacíclica, presente nas excreções eliminadas durante o repasto sanguíneo do inseto vetor. A infecção ocorre devido à entrada da forma infectiva através de descontinuidades na epiderme, por exemplo, a lesão causada pela picada do inseto hematófago, ou ainda por penetração pelas membranas mucosas (BRENER, ANDRADE, BARRAL-NETO, 2000; SOUZA, 2002).

As formas epimastigotas também podem penetrar por esse mesmo processo, porém são destruídas por células fagocíticas, e não influem no restante do processo. Já os tripomastigotas metacíclicos iniciam a invasão das células do hospedeiro, ao se aderirem à superfície celular, sendo então, internalizados através da formação de pseudópodos. Uma vez dentro da célula, os tripomastigotas estão agora envoltos por uma membrana formando o vacúolo parasitóforo.

Este vacúolo irá sofrer fusão com lisossomos da célula hospedeira, acidificando-o, e iniciando o processo de diferenciação da forma tripomastigota para a forma amastigota, ocorrendo também a dissolução da membrana por enzimas secretadas pela forma tripomastigota (BURLEIGH, B.A.& WOOLSEY, 2002). A forma amastigota fica em contato com o citoplasma da célula hospedeira, começando a se dividir por fissão binária. Depois de diversos ciclos de divisão, inicia-se a transformação para as formas tripomastigotas sanguíneas, as quais adquirem um longo flagelo e movimentação intensa, causando a ruptura da célula hospedeira. Após a lise celular, as formas liberadas nas áreas extracelulares podem infectar outras células do organismo ou então serem ingeridas pelo inseto vetor durante o repasto sanguíneo (SOUZA, 2002).

No estômago do inseto vetor inicia-se o processo de diferenciação para a forma epimastigota que, em seguida, migra para o intestino médio e posterior, aderindo-se à superfície, na qual ocorre intensa multiplicação celular. Após esta fase, os epimastigotas sofrem diferenciação para a forma tripomastigota metacíclica (infectante) na parte posterior do aparelho digestivo, mais precisamente no reto, podendo ser eliminados, fechando o ciclo Figura 1.1 (GARCIA & AZAMBUJA, 1991; SOUZA, 2002).

Esta passagem da forma epimastigota para tripomastigota metacíclica é denominada metaciclogênese. Este processo possui um potencial excelente para o estudo dos mecanismos envolvidos na regulação da expressão gênica do parasita (GOLDENBERG, 1990; 1991).

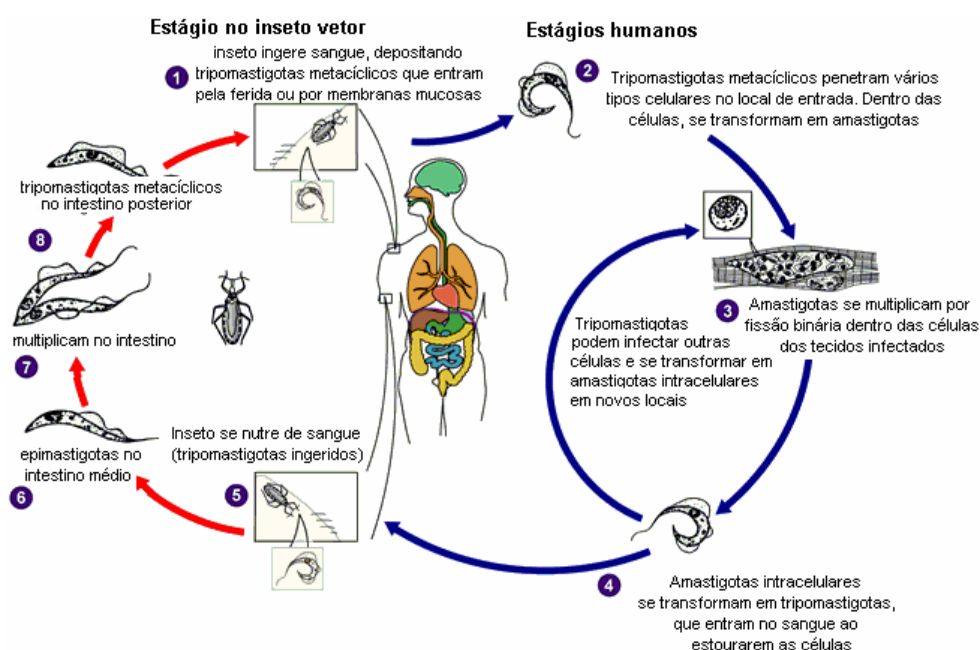


Figura 1.1 Ciclo de vida de *T. cruzi*. (modificado de www.dpd.cdc.gov/dpdx).

De maneira mais detalhada, no sistema digestivo do inseto vetor ocorrem várias alterações químico-físicas que estimulam diretamente modificações na expressão gênica do parasita, o que leva a alterações bioquímicas e morfológicas. Ao serem ingeridos, os parasitas encontram um ambiente com mudanças na osmolaridade e concentração de nutrientes, devido à presença de saliva e enzimas digestivas do estômago do inseto vetor. Além disto, existem fatores hemolíticos que lisam as hemácias e podem afetar o parasita (GOLDENBERG, 1990; KOLLIEN & SCHAUB, 2000). O estômago do inseto tem como função armazenar o sangue obtido, concentrar as proteínas através da remoção da água, armazenar intracelularmente lipídios e lisar a membrana dos eritrócitos. Nesta etapa, diversas enzimas agem e alteram o pH como, por exemplo, fosfatases ácidas, lisozimas e aminopeptidases. Também é nesta porção do tubo digestivo que ocorre a primeira diferenciação, a passagem da forma tripomastigota sanguínea para a forma epimastigota (KOLLIEN & SCHAUB, 2000).

Os epimastigotas migram para o intestino do inseto, o qual possui enzimas que continuam a digestão iniciada no estômago, e nesse local passam por diversos ciclos de divisão celular. Devido à absorção dos nutrientes, a concentração dos mesmos começa a decair e supõe-se que este estresse nutricional influencie na sinalização para o remodelamento da expressão gênica dos parasitas, propiciando o desencadeamento da diferenciação. Na superfície da região posterior do intestino, na camada cuticular do epitélio da glândula retal e do saco retal, ocorre a adesão da forma epimastigota à essa camada, através da interação hidrofóbica do flagelo com as estruturas citadas (ZELEDON *et al*, 1997; KOLLIEN & SCHAUB, 2000). Estas formas aderidas sofrem alongamento e torção, e transformam-se em tripomastigotas metacíclicos.

Há várias evidências de fatores intrínsecos do sistema digestivo do inseto vetor que estimulam o disparo da metaciclogênese em *T. cruzi*, como a composição da urina descarregada no saco retal através dos túbulos de Malpighi (FRAIDENRAICH, *et al*, 1993), um peptídeo resultante da digestão da globina que pode ser capaz de ativar a enzima adenil-ciclase parasitária, presente na forma epimastigota (SCHMIDT *et al*, 1997), AMP cíclico presente na luz do intestino, a diminuição da concentração dos nutrientes devido ao processo de absorção ou a períodos de falta de alimentação do inseto vetor (KOLLIEN & SCHAUB, 2000). No entanto, esses dados são preliminares, se conhecendo relativamente pouco sobre o processo de metaciclogênese.

1.2 A doença de chagas

A doença de Chagas (Chagas, 1909) é característica do continente americano e está atualmente distribuída em 21 países, atingindo México, parte da América Central e toda América do Sul. Há aproximadamente 13 milhões de infectados (TDR,2005), e 120 milhões correm risco de contaminação (www.who.int/tdr/diseases/chagas/swg_chagas.pdf; acesso em 28/11/2007). A doença de Chagas é causada pelo protozoário flagelado *T. cruzi* e a transmissão ao homem pode ocorrer de diferentes formas. A forma mais comum é através de insetos da família Reduviidae que infectam o hospedeiro mamífero durante o repasto sanguíneo, através das fezes contaminadas com o parasita; outras formas, menos comuns, são, por exemplo, por meio de transfusões sanguíneas, transmissão materno-fetal, acidentes laboratoriais e ingestão de formas infectivas.

A doença se caracteriza por duas fases distintas. Na fase aguda pode ocorrer febre, inflamação no local da infecção e nos nodos linfáticos, parasitemia elevada e hepatoesplenomegalia. Porém, normalmente os sintomas não são perceptíveis, devido a sua fraca intensidade. A fase crônica tem evolução lenta com sinais surgindo somente depois de um período de 10 a 20 anos. Durante a fase crônica, podem ocorrer danos irreversíveis ao coração, esôfago ou cólon, os quais apresentam alterações da inervação nervosa e dilatações. A abordagem terapêutica mais utilizada visa o controle da parasitemia com a utilização de quimioterapêuticos, já que não existe profilaxia eficiente por vacinação (COURA & CASTRO, 2002; MARIN-NETO *et al.*, 2002).

Com relação aos aspectos imunológicos básicos, a infecção por *T. cruzi*, como na dos demais parasitas intracelulares, mobiliza múltiplos mecanismos humorais e celulares da resposta imune adaptativa, bem como mecanismos inespecíficos da imunidade inata, os quais levam à eliminação dos parasitas da circulação e a diminuição de sua replicação intracelular. Todavia, o parasita possivelmente persiste indefinidamente, mesmo em níveis mínimos, assim como a resposta imunológica contra o mesmo, resultando em acúmulo de lesões teciduais que acabam por gerar, em alguns casos, alterações morfológicas e funcionais dos tecidos afetados (BRENER, ANDRADE, BARRAL-NETO, 2000). O comportamento do sistema imune sobre os diferentes estágios evolutivos do parasita é distinto. Um exemplo desta grande complexidade de respostas é a ação da via alternativa do complemento: as formas não-infectivas epimastigotas são lisadas por esta via, enquanto a forma infectiva tripomastigota é

resistente à ação lítica do complemento devido à expressão diferencial de proteínas de membrana, como a gp160 e a T-DAF (NOGUEIRA *et al*, 1975; NORRIS *et al*, 1991).

O controle da doença é realizado principalmente através da tentativa de erradicação do inseto vetor de maior importância, o *Triatoma infestans*, através do uso de pesticidas. Porém o fato de ser uma zoonose elimina a possibilidade de erradicação da doença, visto que várias espécies silvestres atuam como reservatório natural e outras espécies de triatomíneos podem transmitir a doença. A partir desta dificuldade, acrescido do alto custo econômico do combate aos insetos, programa este descontinuado em boa parte da América Latina, o entendimento das interações parasita-vetor, como os fatores que induzem o desenvolvimento dos vários estágios do ciclo de vida, podem permitir o desenvolvimento de novas medidas de controle desta parasitose (KOLLIEN & SCHAUB, 2000). O tratamento por quimioterápicos em pacientes infectados visa à eliminação do parasita na forma aguda da doença, porém, como normalmente a infecção passa despercebida, essa estratégia não apresenta resultados epidemiológicos significativos. Portanto, estratégias visando o controle da Doença de Chagas devem focar no bloqueio da infecção ou no tratamento dos pacientes cronicamente afetados.

1.3 Características genômicas de *T. cruzi*

O material genético do *T. cruzi* está localizado em duas estruturas: núcleo e mitocôndria, sendo que o genoma nuclear do *T. cruzi* é estimado entre 40 e 50 Mb, dependendo da cepa estudada (SIMPSON, 1987; ENGLUND *et al*, 1996). Em tripanossomatídeos, não existe condensação cromossômica em nenhum estágio do ciclo celular, e a análise de seu cariótipo necessita de técnicas diferenciadas, como eletroforese em campo pulsado e hibridação com sondas teloméricas (VICKERMAN & TETLEY, 1977). Durante o ciclo celular, a cromatina sofre alteração na sua distribuição, permanecendo concentrada no centro do núcleo, nas formas não replicativas na fase G1, enquanto que na fase S está localizada na periferia nuclear onde ocorre a replicação (ELIAS *et al*, 2002). O cariótipo é estável nas diferentes formas evolutivas do parasita apesar da grande variabilidade intra-específica com relação à vasta variabilidade cromossômica, tanto numérica quanto estrutural (AYMERICH & GOLDENBERG, 1989; HENRIKSSON *et al*, 1995).

Uma das características mais marcantes dos tripanossomatídeos é o fato de que o genoma mitocondrial está organizado em uma complexa rede de moléculas de DNA, localizadas em uma região especializada denominada cinetoplasto. Esta rede é composta por maxicírculos e

minicírculos de DNA. O DNA do cinetoplasto (kDNA) corresponde a cerca de 20 a 25% do DNA total do parasita e é formado por aproximadamente 10.000 a 20.000 minicírculos (1,4 kb em média) variando nas diferentes cepas (MACINA *et al.*, 1985; ENGLUND *et al.*, 1996). Além dos minicírculos, existem também cerca de 20 a 50 maxicírculos (40 kb em média) onde estão presentes genes que codificam proteínas e rRNAs mitocondriais. Os maxicírculos e minicírculos estão relacionados em um processo denominado de editoração que consiste em uma reação que envolve os RNAs mensageiros (mRNA) codificados pelo maxicírculos, bem como a transcrição de RNA guia (gRNA) dos minicírculos. Este processo de editoração consiste na adição ou deleção de bases principalmente uridinas do mRNA para que possam ser corretamente traduzidos (STUART, 1995; HAJDUK *et al.*, 1996; LUKES *et al.*, 2002).

Em 2005 foi publicado o primeiro esboço referente à montagem do genoma de *T. cruzi*. A cepa utilizada para o seqüenciamento foi a CL Brener, por ser uma das mais bem caracterizadas experimentalmente. No entanto, acabou se revelando como tendo um genoma muito complexo, dificultando sua montagem. A cepa CL Brener é híbrida, sendo que dados de diferentes grupos (ZINGALES *et al.*, 1997; BRISSE *et al.*, 1998; MACHADO & AYALA, 2001; GAUNT *et al.*, 2003; BRISSE *et al.*, 2003 EL-SAYED *et al.*, 2005a) são consistentes com o fato de que ela é resultante de dois eventos distintos de hibridização.

O seqüenciamento foi feito através da técnica de *whole genome shotgun* (WGS), sendo que o alto conteúdo de repetições do genoma e a sua natureza híbrida limitaram a eficiência da estratégia de seqüenciamento progressivo de clones em cromossomos artificiais de bactérias. O conteúdo repetitivo do genoma de *T. cruzi* foi inicialmente estimado em 35% (AGUERO *et al.*, 2000), mas acabou revelando-se acima de 50% (EL-SAYED *et al.*, 2005a).

A montagem final do genoma reflete esses problemas, pois contém uma quantidade muito grande de fragmentos genômicos para os quais não se sabe a sua localização cromossômica exata. Estima-se que a cepa CL Brener contenha aproximadamente 64 cromossomos e que o tamanho de seu genoma seja de 87 Mb, com base em análise densitométrica da separação por PFGE (CANO *et al.*, 1995). No entanto, os resultados do projeto de seqüenciamento do genoma estimam que o genoma diplóide esteja entre 106,4 e 110,7 Mb. Desse total, 67 Mb foram selecionados para anotação, consistindo de 5489 *scaffolds*, contendo 8740 contigs, os quais representam o estado extremamente fragmentado da atual versão do genoma. A análise da fração anotada, que perfaz 60,7 Mb, demonstra que 30,5 Mb são compostos de seqüências encontradas pelo menos duas vezes, o que sugere que elas são oriundas dos diferentes haplótipos que compõe o genoma da cepa CL Brener.

Conforme mencionado acima, aproximadamente metade do genoma do *T. cruzi* é composta por seqüências repetitivas como famílias de proteínas de superfície, retrotransposons e elementos subteloméricos. Foram identificados 23.216 elementos codificadores de proteínas dos quais 11.398 possuem função conhecida confirmada experimentalmente ou por homologia, 9.975 genes codificam proteínas hipotéticas conservadas e 1.843 genes codificam proteínas hipotéticas. Além disso, 1.994 genes codificam RNAs estruturais dentre eles: tRNA (RNA transportador) rRNA (RNA ribossomal), snRNA (RNA pequeno nuclear), slRNA (RNA líder), srpRNA (*Signal Recognition Particle*), snoRNA (RNA pequeno nucleolar) (EL-SAYED *et al.*, 2005a). Alguns genes codificadores de proteínas e diversos RNAs funcionais e estruturais no *T. cruzi* mais de duas cópias no genoma diplóide, podendo estar em um único cromossomo ou em diferentes cromossomos. Muitos genes podem apresentar de 4 a 200 repetições, organizados normalmente em *tandem*. Esta organização é vista em diversos genes codificadores de proteínas e RNAs como, por exemplo: proteínas de superfície, ubiquitina, proteínas de choque térmico (HSPs), histonas, calmodulina, tubulinas e amastina (GONZÁLES *et al.*, 1985; DRAGON *et al.*, 1987; MAINGON *et al.*, 1988; REQUENA *et al.*, 1988; SWINDLE *et al.*, 1988; ASLUND *et al.*, 1994; CHUNG *et al.*, 1994; TEIXEIRA *et al.*, 1994; VASQUEZ *et al.*, 1994).

A estimativa do conteúdo gênico é de aproximadamente 12.000, sendo que dos 23.216 modelos gênicos preditos, 6.159 representam alelos presentes no haplótipo IIb, 6.043 representam alelos dos outros haplótipos e 10.368 representam seqüências que não puderam ser atribuídas a um haplótipo (EL-SAYED *et al.*, 2005a).

Com relação à análise de vias biológicas presentes em *T. cruzi* e as suas diferenças com relação a outros organismos, um dos resultados mais impressionantes foi a ausência de diversas classes de proteínas sinalizadoras importantes, como proteínas G heterotriméricas, receptores do tipo serpentina, domínios de interação SH2 e SH3 e fatores de transcrição regulatórios. Por outro lado, *T. cruzi* possui um conjunto muito grande e complexo de proteína-quinases e proteína-fosfatases. Cerca de 167 proteína-quinases distintas foram encontradas em *T. cruzi*, potencialmente ativas, sendo esse número duas vezes maior que o presente em *P. falciparum* (WARD *et al.*, 2004) e 30% maior que o de *S. cerevisiae*. Uma peculiaridade é que a maioria das proteína-quinases de tripanossomatídeos é muito maior do que o necessário para conter um domínio catalítico simples, sugerindo a existência de

domínios acessórios ainda não identificados e que muito de suas interações ainda está por ser descoberta.

Os genomas de outros dois tripanossomatídeos, *T. brucei* (BERRIMAN *et al.*, 2005) e *Leishmania major* (IVENS *et al.*, 2005), também foram publicados recentemente, o que permitiu fazer comparações entre os três patógenos, que compartilham muitas estruturas celulares, mas que possuem diferenças grandes em sua forma de interagir com o hospedeiro. A análise do proteoma compartilhado por esses organismos revelou um conjunto de 6.158 grupos de genes ortólogos (COG, *clusters of orthologous genes*), bem como 1.014 COGs compartilhados por somente dois dos tripanossomatídeos (EL-SAYED *et al.*, 2005b).

Mesmo tendo divergido há mais de 200 milhões de anos (OVERATH *et al.*, 2001; STEVENS *et al.*, 2001), os genomas desses tripanossomatídeos apresentam um grau muito grande de conservação da sintenia. De todos os genes de *T. brucei* e *L. major*, 68 e 75% respectivamente mantém sua sintenia. Além disso, a grande maioria (94%) dos COGs triplos que formam o proteoma central dos três tripanossomatídeos está em regiões de sintenia conservada.

1.4 O *Trypanosoma cruzi* e a transcrição policistrônica

Em eucariotos, éxons e íntrons alternam-se ao longo dos genes. O fenômeno de transcrição é, na maioria dos casos, regulado por uma sequência promotora a montante da região codificadora. Ao transcrito primário, em sua extremidade 5', é adicionado um resíduo metil-guanosina-trifosfato (m7Gppp, ou cap) e, em sua extremidade 3', uma cauda poli-A. Pela reação de *cis-splicing*, são removidos os íntrons deste transcrito e o éxons são ligados (PROUDFOOT *et al.*, 2002).

Os tripanossomatídeos apresentam diferenças com relação a esse perfil. Até recentemente, não havia nenhum indício da existência de íntrons nos genes desses organismos. Essa visão foi modificada em 2000, quando foi identificado um íntron em um gene codificador da enzima poli-A polimerase de *T. cruzi* e *T. brucei* (MAIR *et al.*, 2000). A existência de uma maquinaria de remoção de íntrons foi comprovada pela identificação de um gene codificador do snRNA U1 em *T. brucei* (DIKENGHI *et al.*, 2001), sendo que posteriormente foi possível realizar a identificação do complexo ribonucleoprotéico associado a esse RNA (PALFI *et al.*, 2002). Recentemente, os mesmos pesquisadores ampliaram a descrição desse complexo, incluindo duas proteínas, pouco similares a seus prováveis

homólogos em outros eucariotos, denominadas U1-70K e U1C (PALFI *et al.*, 2005). Embora a maquinaria de *cis-splicing* esteja presente, ela apresenta diferenças estruturais grandes com relação à dos eucariotos superiores, e, embora seja possível a existência de outros exemplos, somente um gene foi identificado como possuindo íntron em tripanossomatídeos até o momento.

Geralmente, os genes em tripanossomatídeos estão distribuídos nos cromossomos de tal forma que é possível ter ilhas de genes codificados na mesma fita de DNA que se estendem por dezenas a centenas de genes (McDONAGH *et al.*, 2000; WORTHEY *et al.*, 2003; MONNERAT *et al.*, 2004; NILSSON & ANDERSSON, 2005; EL-SAYED *et al.*, 2005b). Essa organização é conservada em grande parte dos genomas de *T. brucei* e *L. major* (EL-SAYED *et al.*, 2005b), os quais são os tripanossomatídeos com o genoma mais bem caracterizado, e, possivelmente, o mesmo deve ocorrer para *T. cruzi*. A razão da conservação evolutiva dessa organização ainda não é conhecida.

Normalmente, a distância entre as regiões codificadoras é pequena, o que, em vista da existência de regiões 5' e 3' não-traduzidas (UTR), significa que a região intergênica é muito pequena.

A transcrição nos tripanossomatídeos produz unidades policistrônicas (VANHAMME & PAYS, 1995; TEIXEIRA & DaROCHA, 2003) que são representadas por uma única molécula de RNA geralmente contendo diversos genes que não são relacionados a uma determinada via metabólica, como ocorre nos operons de procariotos. Além disso, os genes de uma mesma unidade de transcrição policistrônica podem apresentar diferenças grandes nos níveis de expressão, demonstrando a importância de mecanismos pós-transcricionais no controle da expressão gênica (VANHAMME & PAYS, 1995). A molécula de RNA policistrônica deve ser processada pelo mecanismo do *trans-splicing* (SUTTON & BOOTHROYD, 1986), a fim de gerar unidades monocistrônicas traduzíveis.

No processo de *trans-splicing*, semelhante ao processo de *cis-splicing*, as unidades codificadoras são separadas e é adicionada na extremidade 5' uma sequência extremamente conservada, espécie-específica, de 39 nucleotídeos, denominada sequência líder (SL) ou mini-éxon (McCARTHY-BURKE *et al.*, 1989; NILSEN, 1992). Nesta reação, participam duas moléculas de RNA, uma doadora e outra aceptora. A doadora é um precursor do mini-éxon, e, em *T. cruzi*, possui cerca de 110 nucleotídeos (ZWIERZYNSKI & BUCK, 1991). A aceptora é o transcrito derivado da unidade policistrônica ao qual o mini-éxon é transferido. O

complexo enzimático envolvido nesta reação é semelhante ao descrito para a de *cis-splicing* (LAIRD, 1989).

Os transcritos processados são estabilizados pela estrutura cap 4 do mini-éxon (ULLU & TSCHUDI, 1991) e pela adição de uma cauda poli-A, à extremidade 3', sendo que a adição de ambos os elementos ocorrem em conjunto (LEBOWITZ *et al.*, 1993). Regiões ricas em resíduos de pirimidinas, além da sequência consenso (AG) para o *trans-splicing*, regulam a adição do mini-éxon e da cauda poli-A. Deleções nestas regiões impedem o correto processamento destes transcritos (MATTHEWS *et al.*, 1994).

Os três tipos de RNA polimerase, I, II e III, descritos em eucariotos superiores, já foram observados em tripanossomatídeos (CORNELISSEN *et al.*, 1989; CORNELISSEN *et al.*, 1990; KOCK & CORNELISSEN, 1991). Os genes que codificam para as suas subunidades principais foram clonados e caracterizados nestes protozoários (EVERS *et al.*, 1989; JESS *et al.*, 1989; KOCK *et al.*, 1988; SMITH *et al.*, 1989). Embora tenham sido encontradas poucas diferenças estruturais em relação às RNA polimerases de eucariotos superiores, destacando-se a modificação da extensão terminal da RNA polimerase do tipo II, esta comporta-se de maneira diferente em presença de inibidores específicos como, por exemplo, a α -amanitina. Do mesmo modo, apresentam diferenças na dependência de Mn^{++} , na utilização de fita simples e no reconhecimento de promotores específicos. Em *T. cruzi*, apenas promotores associados a genes ribossomais e ao mini-éxon foram caracterizados (DIETRICH *et al.*, 1993; NUNES *et al.*, 1997). A atividade destes promotores de *T. cruzi* parece ser controlada por fatores de transcrição que são cepa-específicos. Assim, o promotor ribossomal da cepa CL não é funcional em outras cepas (TYLER-CROSS *et al.*, 1995; NUNES *et al.*, 1997). Em outros tripanossomatídeos, promotores equivalentes e fatores de transcrição associados foram descritos e caracterizados (DAS & BELLOFATTO, 2003; LANDFEAR, 2003).

Nestes parasitas, de modo geral, a RNA polimerase I transcreve os genes ribossomais e a RNA polimerase III transcreve moléculas de RNA nucleolar e de transferência. Os genes do mini-éxon são transcritos pela RNA polimerase II (GILINGER & BELLOFATTO, 2001), a qual também transcreve os mRNAs. Contudo, as polimerases e os promotores associados à transcrição dos genes que codificam proteínas ainda não foram identificados (LANDFEAR, 2003).

1.5 A regulação da expressão gênica em *T. cruzi*

Na maioria dos eucariotos, a regulação da expressão gênica é, em grande parte, realizada no início da transcrição pela RNA polimerase II, sendo esta a enzima responsável pela transcrição de genes codificadores de proteínas e genes de pequenos RNAs nucleares envolvidos no processamento de outros RNAs. Os rRNAs 18S e 28S são transcritos pela RNA polimerase I e a RNA polimerase III transcreve genes codificadores de tRNA, rRNA 5S e pequenos RNAs nucleares e citoplasmáticos. No *T. cruzi*, a RNA polimerase I é responsável pela transcrição de alguns genes codificadores de proteínas sendo esta a principal diferença quando comparadas a outros eucariotos. Outra grande diferença é que aparentemente não existe em tripanossomatídeos regulação por promotores específicos para polimerase II. Há evidências de que a transcrição da maioria dos genes codificadores de proteínas é ubíqua (VANHAMME & PAYS, 1995; CLAYTON, 2002). Com relação aos promotores, em tripanossomatídeos não há evidências de sua existência para os genes codificadores de proteínas, tendo sido identificados somente para os genes do mini-éxon. Assim, os mecanismos de regulação da expressão gênica possivelmente ocorrem etapas posteriores à transcrição, no que é denominado de regulação pós-transcricional.

Conforme descrito acima, a expressão gênica em *T. cruzi* apresenta diversas características peculiares, como a transcrição policistrônica, o processo de *trans-splicing* (editoração do RNA mensageiro), a edição do RNA mitocondrial e a transcrição dos genes codantes de proteínas pela RNA polimerase do tipo I e a ausência de seqüências promotoras típicas (TEIXEIRA & daROCHA, 2003). Todos esses elementos reforçam a idéia de que a regulação da expressão gênica em tripanossomatídeos possui características distintas com relação à de outros organismos. O fato de que genes presentes na mesma unidade policistrônica apresentam níveis de mRNA processado distintos reforça a idéia de que a regulação seja ao nível pós-transcricional. Estas características específicas dos tripanossomatídeos juntamente com a relevância médica o tornam um bom alvo para estudo visando a melhor compreensão de sua biologia como, por exemplo, os mecanismos de controles transcricionais e pós-transcricionais.

As formas epimastigotas e amastigotas, quando iniciam o processo de diferenciação para as formas não replicativas e infectivas, apresentam uma grande diminuição na atividade da RNA polimerase I e II, porém a redução no nível de transcrição é geral e não gene dependente (ELIAS *et al.*, 2001). Deste modo, é possível observar que a expressão de alguns genes pode

manter-se constante ou aumentar, mesmo com a queda dos níveis de transcrição (ABUIN *et al.*, 1999; TOMAS & KELLY, 1996; RECINOS *et al.*, 2001). Também é possível observar proteínas de expressão transiente nas formas em diferenciação (CONTRERAS *et al.*, 1985). Em *T. cruzi*, foi demonstrado que moléculas de RNA mensageiro podem ser mantidas estáveis e não associadas a polissomos, desmobilizadas no citoplasma, sendo passíveis de tradução *in vitro* (GOLDENBERG *et al.*, 1985). Há evidências de que o controle da associação à maquinaria de tradução seja o mecanismo principal na expressão de alguns genes diferencialmente expressos durante a metaciclogênese de *T. cruzi* (ÁVILA *et al.*, 2001; DALLAGIOVANNA *et al.*, 2001; FRAGOSO *et al.*, 2003).

Além disso, experimentos realizados com o gene da amastina de *T. cruzi* e *Leishmania* indicam que a sua expressão é dependente de seqüências contidas em sua região 3'-UTR. Tais seqüências seriam responsáveis pela associação desta molécula de RNA mensageiro aos polissomos e pelo aumento de sua taxa de tradução, sem que a estabilidade deste transcrito fosse alterada (TEIXEIRA *et al.*, 1994; TEIXEIRA *et al.*, 1995; COUGHLIN *et al.*, 2000; BOUCHER *et al.*, 2002).

A relevância da porção 3'-UTR na regulação da expressão gênica também pode ser avaliada em experimentos de transfecção envolvendo genes estágio-específicos (gp85 em tripomastigotas sanguíneos e amastigotas, gp72 em epimastigotas, gp82 em tripomastigotas metacíclicos) ou de expressão constitutiva (gGAPDH e hsp60) e o gene da luciferase, como gene repórter. Nestes ensaios, a região 3'-UTR dos genes em estudo foram inseridas logo depois do gene da luciferase e, após a transfecção das diferentes formas evolutivas de *T. cruzi* com estas construções, foi possível observar que construções com as regiões 3'-UTR específicas justapostas ao gene da luciferase tiveram padrão de expressão idêntico aos de seus genes de origem (NOZAKI & CROSS, 1995).

Por outro lado, há descrições de que a estabilidade dos transcritos de certos genes em tripanossomatídeos varia ao longo do ciclo celular, alterando a sua expressão (HEHL *et al.*, 1994; BERBEROF *et al.*, 1995; BLATTNER & CLAYTON, 1995; NOZAKI & CROSS, 1995; TEIXEIRA *et al.*, 1995). Muitos fatores que contribuem para esta estabilidade foram caracterizados (CLAYTON, 2002; D'ORSO *et al.*, 2003; SBICEGO *et al.*, 2003).

Seqüências presentes nas regiões intergênicas e 5'-UTR também podem estar relacionadas à regulação da expressão gênica em *T. cruzi*. Recentemente, foi descrita uma proteína que reconhece especificamente motivos poli [dT-dG] em fitas simples de

polinucleotídeos, os quais são muito freqüentes nas regiões intergênicas de *T. cruzi*, e que pode estar envolvida em mecanismos de regulação gênica (DUHAGON *et al.*, 2003).

1.6 A metaciclogênese

Durante o seu ciclo evolutivo o *T. cruzi* alterna entre diferentes tipos morfológicos com propriedades características, incluindo a infectividade para o hospedeiro vertebrado (DE SOUZA, 1984). Assim, o estudo do processo de diferenciação, além de permitir a melhor compreensão de um fenômeno biológico básico, pode levar à identificação de alvos e processos com potencialidade de contribuir na busca de agentes profiláticos para a doença de Chagas. Nesse sentido, avanços significativos foram alcançados na tentativa de simular em laboratório as condições naturais que podem levar o parasita a se diferenciar, possibilitando o estudo da reprogramação gênica a que estes organismos são submetidos durante o processo de diferenciação, sobretudo para as formas infectivas.

Uma etapa crucial do ciclo de vida do *T. cruzi* é a metaciclogênese, onde formas não infectivas, epimastigotas, diferenciam-se em formas infectivas, tripomastigotas metacíclicas (DE SOUZA, 1984). Esse processo é um modelo apropriado para o estudo dos mecanismos envolvidos na regulação da diferenciação e da expressão gênica no parasito. Uma etapa essencial para o aprofundamento do estudo desse processo biológico foi o desenvolvimento de um meio com condições quimicamente definidas que mimetizam o processo que ocorre no inseto vetor da doença de Chagas (CONTRERAS *et al.*, 1985; BONALDO *et al.*, 1988), usando-se como padrão a cepa Dm28c de *T. cruzi* (CONTRERAS *et al.*, 1988). A compreensão desta etapa do ciclo do parasita sempre foi objeto de grande interesse pelo potencial de contribuir para a elucidação dos mecanismos que modulam a expressão de genes estágio-específicos, um processo de diferenciação e o estabelecimento da infectividade do parasita.

A diferenciação de epimastigotas a tripomastigotas metacíclicos requer a adesão dos epimastigotas a um substrato (BONALDO *et al.*, 1988), semelhante ao que foi descrito para o processo de metaciclogênese no interior do triatomíneo (SCHAUB, 1988). Os fenômenos de adesão e de diferenciação celular são interrompidos se, após um período de estresse nutricional, as condições de nutrição forem restauradas. Deste modo, é possível que o estresse nutricional esteja associado à expressão de proteínas de adesão (FIGUEIREDO *et al.*, 2000).

De fato, a metaciclologênese no inseto vetor ocorre na porção terminal do intestino do triatomíneo onde ocorre a diminuição dos nutrientes devido a absorção dos mesmos pelo inseto, bem como a utilização dos nutrientes pelos parasitas. Além disso, períodos sem ocorrência de alimentação ocorrem ocasionando o estresse nutricional (KOLLIEN & SCHAUB, 2000). Estudos de metaciclologênese *in vitro* demonstram a importância do estresse nutricional no disparo da diferenciação (CONTRERAS *et al.*, 1985). A escassez de nutrientes está associada ao processo de acidificação do conteúdo do reservossomo e de ativação de suas enzimas, com a degradação das proteínas acumuladas nessa organela, o que leva ao seu desaparecimento, com a liberação dos aminoácidos para o citoplasma (SOARES *et al.*, 1989; URBINA, 1994; SOARES, 1999).

É razoável presumir que durante o processo de adesão, que é necessário porém não suficiente para a diferenciação (BONALDO *et al.*, 1988; KLEFFMANN *et al.*, 1998), os parasitas tenham a sua expressão gênica alterada a partir de sinais externos oriundos do processo de adesão (GOLDENBERG *et al.*, 1985). Outros fatores também foram descritos como capazes de induzir o processo de metaciclologênese, tais como fatores séricos (CONTRERAS *et al.*, 1985, DUSANIC, 1980) e estresse nutricional (CASTELLANI *et al.*, 1967; CONTRERAS *et al.*, 1985). Um fator importante como desencadeador da metaciclologênese do *T. cruzi* é o AMP cíclico (cAMP) (GONZALES-PERDOMO *et al.*, 1988), já que aumentos na concentração intracelular deste elemento levam à transformação de epimastigotas em tripomastigotas metacíclicas. O mecanismo através do qual o cAMP exerce o seu efeito permanece indeterminado. Todavia é interessante notar que foram observadas alterações nos níveis de expressão de proteína-quinases (ULLOA *et al.*, 1988; OCHATT *et al.*, 1993), indicando o envolvimento de mensageiros secundários e a transdução de sinal durante o processo.

No caso da metaciclologênese no inseto vetor, o cAMP é excretado pelo túbulos de Malpighi (KOLLIEN & SCHAUB, 2000) e a atividade de adenilato ciclase é estimulada por hemoglobina, normalmente encontrada no intestino do inseto vetor (FRAIDENRAICH *et al.*, 1993). Interessantemente, peptídeos sintéticos correspondentes à sequência de alfa-globina foram capazes de estimular a metaciclologênese quando adicionados à dieta dos triatomíneos (GARCIA *et al.*, 1995).

Outra alteração metabólica observada durante a metaciclologênese refere-se ao acúmulo de lipídios e à mudança de composição lipídica do parasita, havendo redução dos níveis dos ácidos mirístico e linoléico e aumento das concentrações dos ácidos palmítico, esteárico e

linolênico (ESTEVEZ *et al.*, 1989). Logo após serem submetidas ao estresse nutricional *in vitro*, as células em diferenciação são mais semelhantes às tripomastigotas do que às epimastigotas, considerando o seu conteúdo lipídico (SOARES *et al.*, 1989; ESTEVES *et al.*, 1989). Estas mudanças são também acompanhadas por modificações nos açúcares de superfície (DE ANDRADE *et al.*, 1991), acarretando em modificações na permeabilidade e fluidez de membrana dos parasitos.

O processo de metaciclógênese também acarreta modificações no perfil de expressão gênica dos parasitos, e as alterações no padrão de genes expressos podem preceder as alterações morfológicas que ocorrem no processo de diferenciação (CONTRERAS *et al.*, 1985). Os níveis de transcrição são reduzidos durante o processo (GOLDENBERG *et al.*, 1985) e vários genes são transitoriamente expressos, com um perfil de regulação de sua expressão nitidamente pós-transcricional, sendo que alguns deles foram caracterizados como tendo seu mecanismo de regulação através de sua mobilização diferencial para os polissomos (ÁVILA *et al.*, 2001; DALLAGIOVANA *et al.*, 2001; FRAGOSO *et al.*, 2003). Seis horas depois do início do processo de diferenciação, as formas epimastigotas em diferenciação apresentam-se resistentes à lise pelo complemento, uma propriedade das formas tripomastigotas metacíclicas (NOGUEIRA *et al.*, 1975).

Outros fatores também foram descritos como capazes de afetar o processo de diferenciação. Assim, inibidores de poli-ADP ribosilação (ISOLA *et al.*, 1987) ou inibidores de proteases (BONALDO *et al.*, 1991) inibem o processo de diferenciação enquanto a L-prolina estimula a diferenciação (CONTRERAS *et al.*, 1985; HOMSY *et al.*, 1989). É interessante ressaltar que a L-prolina também foi descrita como um fator de estimulação da diferenciação de amastigotas em tripomastigotas no interior das células do hospedeiro vertebrado (TONELLI *et al.*, 2004).

A presença e a concentração dos íons Ca^{++} , K^+ , PO_4^- e HCO_3^- facilitam a metaciclógênese, enquanto o ânion Cl^- a impede (KRASSNER *et al.*, 1991). As concentrações de polifosfatos inorgânicos de cadeia longa e curta, nos acidocalcissomos do *T. cruzi*, elevam-se durante a amastigogênese e na fase estacionária de crescimento das formas epimastigotas. Estas concentrações são reduzidas em condições de estresse hiposmótico ou alcalino. A hidrólise destas cadeias está associada à liberação de Ca^{++} pelos acidocalcissomos e a respostas rápidas do parasita às alterações ambientais. Portanto, os polifosfatos também podem estar envolvidos na regulação da metaciclógênese (RUIZ *et al.*, 2001).

A indução da metaciclo gênese em meio quimicamente definido permite isolar células em vários estágios de diferenciação. *In vitro*, a transferência de epimastigotas de um meio nutricional rico para um meio sem qualquer fonte de carbono gera o estresse nutricional necessário para o começo do processo (CONTRERAS *et al.*, 1985; BONALDO *et al.*, 1988; CONTRERAS *et al.*, 1988). Há duas décadas, a metaciclo gênese *in vitro* tem sido o modelo principal do nosso grupo, o qual está sendo utilizado para caracterizar a expressão gênica nas diferentes etapas da metaciclo gênese em *T. cruzi* e estudar os seus mecanismos de regulação da expressão gênica (ÁVILA *et al.*, 2003).

1.7 O estudo genômico funcional de *T. cruzi*

A partir de meados da década de 80, com a iniciativa do projeto genoma, inúmeros organismos foram alvos de seqüenciamento genômico. Muitos consórcios para mapeamento e seqüenciamento foram estabelecidos, e vários destes projetos já foram finalizados e diversos outros estão em andamento. Com relação ao *T. cruzi*, o seqüenciamento já foi terminado e publicado o primeiro esboço referente à montagem do seqüenciamento, conforme relatado anteriormente. O próximo desafio após esta etapa consiste em compreender como os conjuntos gênicos interagem através da regulação da expressão do genoma em diferentes situações de interação do organismo com o ambiente e outros organismos. O principal objetivo é compreender o funcionamento celular, visando o desenvolvimento de novas abordagens terapêuticas, como vacinas e drogas, e melhoramentos no campo da agropecuária e biotecnologia. O conhecimento apenas das seqüências gênicas e suas funções não garante a compreensão dos processos acima citados, sendo necessário uma abordagem ampla, através de inúmeras ferramentas que estão surgindo e sendo melhoradas constantemente (DHAND, 2000; LOCKHART & WINZELER, 2000; DEGRAVE, 2001; BLAXTER & IVENS, 1999; YANG & SPEED, 2002). Visto esta necessidade, o Instituto de Biologia Molecular do Paraná (IBMP) investiu em uma abordagem genômica funcional, primeiramente com a implantação da plataforma de microarranjo de DNA bem como investimentos na área de bioinformática.

A técnica de microarranjo foi descrita inicialmente por SCHENA *et al.* (1995), em um artigo descrevendo a aplicabilidade da técnica na quantificação de um conjunto restrito de genes de *Arabidopsis thaliana*. Na sua forma mais geral, um arranjo de DNA é usualmente um substrato, uma membrana de náilon ou nitrocelulose, vidro ou plástico, sobre o qual são depositados diversos DNAs fixa simples distintos com seqüências diferentes. Usualmente,

esse DNA é colocado de forma a criar estruturas localizadas que são arranjadas em um padrão regular, semelhante a uma grade, dando-se o nome de sonda ao material que é depositado na superfície do substrato. Se a quantidade de sondas distintas é grande e são utilizados sistemas de produção robótica, as quais permitem a criação de uma densidade muito alta de sondas em uma região muito pequena do substrato, dá-se o nome de microarranjo de DNA ao arranjo de DNA resultante.

O material a ser colocado no microarranjo pode ser de diferentes tipos, de acordo com o propósito do mesmo. Atualmente, os três tipos mais comumente encontrados de microarranjo são os que contém como sonda produtos de PCR dupla fita (geralmente, oriundos de bibliotecas de cDNA), oligonucleotídeos fita simples longos (entre 45 e 70 mer) e oligonucleotídeos fita simples curtos (geralmente menores que 30 mer). Essas sondas geralmente são complementares a regiões do genoma que codificam proteínas e, portanto, os microarranjos são principalmente focados para a avaliação do transcriptoma codificador de proteínas, embora qualquer tipo de avaliação por hibridação de ácidos nucleicos possa ser implementado.

O microarranjo é colocado em contato com uma solução contendo a população de mRNA que desejamos estudar, a qual é denominada normalmente de alvo, e essa preparação é deixada sobre controle rigoroso de temperatura para que o pareamento das regiões complementares na sonda e no alvo, que estão ambos em fita simples, possa ocorrer, no fenômeno denominado hibridação. Após ocorrer a hibridação, as moléculas do alvo que não estão pareadas ou que apresentem pareamento fraco são retiradas através de lavagens sucessivas, que afetam fracamente a ligação dos alvos corretamente pareados com a sonda. Finalmente, a ocorrência da hibridação é passível de ser quantificada, pois as moléculas da população alvo foram marcadas com a adição de agentes fluorescentes ou radioativos, cujo sinal pode ser captado por dispositivos especializados.

Recentemente, PROBST (2005) avaliou o transcriptoma de *T. cruzi* durante o processo de metaciclogênese através do uso de um microarranjo de DNA contendo aproximadamente 6.200 sondas, desenvolvido no IBMP. Cerca de 1.500 genes foram identificados como diferencialmente expressos durante esse processo, sendo que aproximadamente 700 são proteínas hipotéticas.

Estes resultados constituem a primeira avaliação em larga escala da regulação dos mecanismos de transcrição e tradução dos mRNAs. A complementação dessa análise com outras abordagens relativas a diferentes níveis de organização biológica, como caracterização

protéica, interação entre proteínas, entre outros, permitirá a obtenção de resultados mais robustos e com maior impacto no entendimento da biologia molecular do *T. cruzi*.

Dentre os métodos de análise em larga escala que complementam os resultados de microarranjos, um dos mais diretamente relacionados e que pode ser feito, atualmente, de maneira relativamente fácil, é a proteômica ou estudo do proteoma. Primeiramente o termo proteoma foi utilizado para descrever conjuntos de proteínas codificadas pelo genoma (WILKINS *et al*, 1996). Porém, atualmente a proteômica engloba não só a descoberta das proteínas de uma célula, mas também a investigação de isoformas, modificações protéicas, interações entre elas, descrição estrutural, construções de modelos de vias ou mapas de interações (EISENBERG *et al*, 2000; PATTERSON & AEBERSOLD, 2003; TYERS & MANN, 2003), entre outros.

A proteômica constitui-se em uma rica fonte de informação biológica decorrente do envolvimento das proteínas em quase todas atividades biológicas, acrescido das diversas propriedades que combinadas geram sistemas de grande complexidade. Baseado nisto, a proteômica pode propiciar o entendimento dos mais variados sistemas biológicos. Esta abordagem, bem como as outras abordagens referentes a genômica funcional, são consideradas como sendo uma *discovery science*, ou seja, uma ciência baseada em descobertas e não em hipóteses, com um enfoque mais investigativo, normalmente fundamentadas na geração e obtenção de um grande número de dados, para permitir uma visão mais ampla e completa dos processos. Por exemplo, o número de resultados provenientes dos estudos de microarranjo gerou uma base de dados muito grande, permitindo, juntamente com abordagens complementares, sugerir ou comprovar novas hipóteses. Em resumo, com este grande número de resultados obtidos, pode-se integrar os dados, formular teorias e caracterizar os mais variados sistemas (PATTERSON & AEBERSOLD, 2003). Ao mesmo tempo, o excesso de resultados gerados bem como suas interrelações, podem gerar dificuldades na comprovação biológica devido a grande complexidade.

A partir desta nova perspectiva, o objetivo principal é determinar como, quando e sobre quais condições as proteínas foram geradas, modificadas, agregadas a complexos e degradadas. Embasada por diversas técnicas de estudo, a proteômica abriu um campo sem precedentes para pesquisas e seu uso é extremamente extenso. Muitos trabalhos já foram publicados e é difícil se ter um panorama da sua aplicação, devido a sua extensão. Por exemplo, em humanos já existem trabalhos mostrando resultados da análise proteômica do nucléolo (ANDERSEN *et al.*, 2002), da resposta ao interferon por células do fígado (YAN *et*

al., 2004), de estudo de nefropatia causada por diabete (THONGBOONKERD *et al.*, 2004, análise proteômica de câncer de ovário (WANG *et al.*, 2004), entre outros.

Com relação à proteômica direcionada aos estudos de parasitas, existem também inúmeras publicações com diversos parasitas e abordagens, como por exemplo, o estudo de novos alvos terapêuticos e da resistência de *Plasmodium falciparum* (COOPER & CARUCCI, 2004), complexo de adesão do *Toxoplasma gondii* (ZHOU *et al.*, 2004) e diferenciação de estágio evolutivo de *Leishmania mexicana* (NUGENT *et al.*, 2004). No caso do *T. cruzi*, a avaliação proteômica das formas do ciclo evolutivo já foi publicada, evidenciando-se algumas proteínas diferencialmente expressas (PABA *et al.*, 2003, PABA *et al.*, 2004; PARODI-TALICE *et al.*, 2004; ATWOOD *et al.*, 2005).

Baseados na necessidade de obtenção de mais dados para serem agregados aos dados provenientes da genômica e proteômica, visando a melhor compreensão dos processos celulares e moleculares, diversas outras abordagens estão surgindo como, por exemplo, metaboloma, ORFeoma, ribonoma, interatoma, entre outros. Todas elas geram um número ainda maior de dados permitindo uma visão mais completa, ou sistêmica, dos processos biológicos, dentro de uma nova área de pesquisa denominada biologia de sistemas (*system biology*).

1.8 Proteínas hipotéticas em *T. cruzi*

Com a liberação dos dados do seqüenciamento do genoma, foi possível fazer uma análise utilizando-se ferramentas de predição gênica para avaliar a classificação funcional de proteínas. No caso do *T. cruzi*, sendo um organismo relativamente pouco estudado e distante evolutivamente dos modelos biológicos mais clássicos, muitas regiões codificadoras preditas não tiveram uma anotação funcional confiável e foram classificadas genericamente como proteínas hipotéticas. Esse conjunto de genes, constituindo 52,3% das 22.570 proteínas identificadas, é dividido em dois grupos, denominados hipotéticas conservadas, as quais apresentam ortólogos em outros organismos, ou somente hipotéticas que, a princípio, seriam restritas a *T. cruzi*. A proporção de proteínas hipotéticas de *T. cruzi* é alta, mas não difere muito de outros organismos, mesmo aqueles melhor estudados.

Quando analisamos as formas infectivas, tripomastigotas e tripomastigotas metacíclicas por microarranjo, observamos que uma grande proporção dos genes aumentados nessas fases é composta por codificadores de proteínas hipotéticas. Portanto, para um melhor

entendimento do processo de metaciclogênese, é essencial o esclarecimento funcional das proteínas codificadas por esses genes, por seu interesse biológico e pelo potencial como alvo para o desenvolvimento de novas drogas antiparasitárias.

2 Objetivos

2.1 *Objetivo geral*

Caracterizar genes codificadores de proteínas hipotéticas diferencialmente expressas em tripomastigotas metacíclicos, selecionados a partir de dados provenientes da análise do ciclo de vida de *T. cruzi* obtidos através da metodologia de microarranjo de DNA.

2.2 *Objetivos Específicos*

- Avaliar o perfil das proteínas hipotéticas diferencialmente expressas durante a metaciclogênese;
- Selecionar genes codificadores de proteínas hipotéticas, baseado no perfil de expressão diferencial;
- Avaliar características bioinformáticas das proteínas hipotéticas identificadas como diferencialmente expressas durante o processo de metaciclogênese de *T. cruzi*;
- Clonar os genes selecionados utilizando a metodologia Gateway®;
- Expressar os genes selecionados em *E. coli*;
- Obter anticorpos policlonais a partir das proteínas recombinantes expressas;
- Realizar ensaios de *Western blot* utilizando os soros obtidos;
- Analisar os dados de *Western blot* pela extração de intensidade a fim de verificar variações durante a metaciclogênese e sua correspondência com as mudanças do transcriptoma;
- Realizar imunolocalização por microscopia óptica a partir dos anticorpos policlonais obtidos.

3 Justificativas

O protozoário flagelado *T. cruzi* é o causador da doença de Chagas, que atualmente, atinge 17 milhões de pessoas em 21 países do continente americano, com 120 milhões de pessoas correndo o risco de contaminação. Apesar de ser estudado há mais de noventa anos sobre os diferentes aspectos envolvidos em sua patologia, sendo um dos parasitas mais estudados, não existem vacinas ou medicamentos efetivos para o tratamento da doença. Na escala evolutiva, o *T. cruzi* é classificado como um dos organismos eucarióticos mais antigos. Algumas de suas características genômicas são peculiares, como por exemplo, mecanismos singulares de expressão gênica, *trans-splicing* e editoração do RNA.

Inserido no ciclo evolutivo do parasita, o processo de diferenciação celular da metaciclogênese é uma etapa importante onde ocorrem transformações estruturais e bioquímicas já conhecidas. Sabendo-se que a diferenciação celular se dá com a transformação de uma forma não infectante (epimastigota) para uma forma infectante (tripomastigota metacíclico), é de extrema importância a compreensão dos mecanismos envolvidos com a regulação da expressão gênica do parasita nesta fase do ciclo evolutivo.

Além disto, a velocidade com que novos genomas têm sido seqüenciados faz com que o acúmulo de dados de seqüências de DNA não possa ser acompanhado no mesmo ritmo pela elucidação da funcionalidade dos elementos neles codificados. A técnica de microarranjo representa uma abordagem que visa, de forma inicial, categorizar a importância funcional dessa informação.

Através de análise de clusterização hierárquica, por exemplo, é possível fazer inferências sobre a função gênica. Entretanto, um grande número dos genes seqüenciados permanece como codificadores de proteínas cuja função, em maior ou menor grau, é desconhecida. Baseado neste panorama, a escolha de grupos gênicos co-regulados a partir dos dados obtidos pela metodologia de microarranjo e programas de clusterização permite uma escolha bastante robusta para a seleção de genes candidatos para a clonagem e caracterização funcional, agregando assim um maior conhecimento acerca dos genes e possibilitando novas inferências sobre as vias biológicas e suas interações.

Sendo assim, juntando esforços ao aumento do conhecimento referente ao *T. cruzi*, pode-se salientar a necessidade da produção do ORFeoma, abordagem esta que permitirá agregar outros aspectos com os dados de genômica e proteômica, já que o IBMP está desenvolvendo vetores para o próprio *T. cruzi*, visando análises de co-localização, super-expressão, interação

e formação de complexos protéicos. Neste sentido, visando à implantação posterior de uma plataforma mais robusta para a construção de uma biblioteca de clones, o presente trabalho também visa identificar o conhecimento necessário para adquirir experiência de formação nas técnicas utilizadas, juntamente com o desejo de se obter dados para comprovação de resultados provenientes do microarranjo de DNA e proteômica desenvolvidos no IBMP.

Justificando e exemplificando a necessidade da produção do ORFeoma de *T. cruzi*, os vetores que estão sendo desenvolvidos podem gerar dados, por exemplo, de localização e co-localização celular, baseado em tecnologias de clonagens que permitem ingressar em plataformas *high-throughput* com grande facilidade (RUAL *et al*, 2004). Os dados de localização celular podem fornecer vários indícios de função, estado de ativação, interação com outras moléculas, fatores estes que podem indicar eventos de sinalização celular, progressão do ciclo celular e mudanças ambientais (O'ROURKE *et al*, 2005).

4 Material e Métodos

A análise do processo de metaciclo gênese pelo uso de microarranjos de DNA, através da perspectiva de genômica funcional, foi realizada anteriormente no IBMP em uma tese de doutoramento (PROBST, 2005). Devido à relevância da estruturação e resultados desse trabalho para o desenvolvimento da presente dissertação, optou-se por fazer uma descrição sucinta das etapas e informações essenciais contidas no mesmo, sendo importante reforçar que essas etapas laboratoriais não foram realizadas no presente trabalho, mas por haver uma forte ligação entre os mesmos, no sentido de continuidade, estão descritas em material e métodos. O presente trabalho teve como ponto de partida o item 4.4 utilizando os resultados anteriores de PROBST, 2005.

4.1 Confecção dos microarranjos de DNA

O nosso grupo trabalha com microarranjos desde 2001, quando saíram os primeiros resultados preliminares utilizando esta metodologia. Desde 2003, possuímos a capacidade plena do processo de produção e toda a manufatura e produção dos resultados, bem como análise do microarranjo é feita no IBMP.

O microarranjo de *Trypanosoma cruzi* se encontra atualmente na sua versão 5.1 possuindo uma cobertura aproximada de 7000 genes.

4.2 Arquitetura do microarranjo

As sondas, compostas a partir de bibliotecas de EST de *T. cruzi*, em um total de 6.198 elementos, estão presentes em cada lâmina em três réplicas, cada uma delas contendo 16 setores de 20 colunas e 20 linhas, num total de 6.400 posições disponíveis, sendo que as posições restantes (n=202) estão ocupadas por solução de espotagem (DMSO 50%), como controle de fluorescência basal, ou vazias. Os 6.198 spots podem ser divididos em:

- 3.552 sondas individuais, amplificadas de DNA genômico de *T. cruzi*, oriundas da segunda versão do microarranjo, obtidas a partir dos dados de EST disponíveis no Genbank no ano de 2000.
- 2.354 sondas individuais, amplificadas de insertos em vetor Topo T-A (Invitrogen), referentes à biblioteca de EST de epimastigotas, epimastigotas em estresse nutricional,

epimastigotas aderidos em 24 horas de diferenciação e tripomastigotas metacíclicos, produzida no IBMP (PICCHI & PROBST, em preparação).

- 144 sondas referentes a 12 genes selecionados como controles de hibridação, por apresentarem uma intensidade de sinal muito forte nos primeiros experimentos feitos. (n= 8 a 20 sondas por gene, em número variável).
- 134 sondas referentes a genes selecionados, incluídos por serem de interesse às linhas de pesquisa de outros grupos do IBMP ou colaboradores.
- sondas referentes ao genoma do endossimbionte de *Crithidia deanei*, utilizados como controles negativos de DNA.
- 10 produtos da amplificação do gene Q do bacteriófago λ Q, usado como controle negativo ou positivo externo, caso adicionado o mRNA do gene à preparação do material a ser hibridado.

Devido à complexidade e redundância da montagem atual do genoma de *T. cruzi*, é possível fazer somente uma estimativa dos genes distintos que são reconhecidos por essas sondas. Nossas estimativas atuais demonstram que um número mínimo de 4.000 e um máximo de 5.000 genes estão sendo avaliados pela versão atual do microarranjo de *T. cruzi*.

4.3 Desenho experimental

Foram realizados dois experimentos distintos de metaciclogênese, denominados Met16 e Met19. Em cada experimento, foram obtidas as amostras epimastigotas (Epi), epimastigotas em estresse nutricional (Str), epimastigotas aderidos e em diferenciação por 3, 12 e 24 horas (3h, 12h e 24h), das quais foram extraídos mRNA total e polissomal da mesma amostra biológica.

Todos os pontos de cada metaciclogênese são oriundos do mesmo conjunto inicial de epimastigotas em fase logarítmica de crescimento (Epi), para que as diferenças biológicas inerentes a cada réplica biológica de metaciclogênese pudessem ser homogeneizadas internamente.

As amostras de tripomastigotas metacíclicos foram obtidas separadamente das metaciclogêneses descritas acima, constituindo-se em duas réplicas biológicas, as quais foram comparadas com os epimastigotas de Met16 e Met19.

Cada metaciclôgênese foi hibridada com microarranjos oriundos do mesmo lote de fabricação, sendo utilizados três lotes distintos. As hibridações entre epimastigotas e tripomastigotas metacíclicas foram realizadas com microarranjos provenientes de um quarto lote.

Os diagramas representativos das hibridações realizadas podem ser vistos na Figura 4.1.

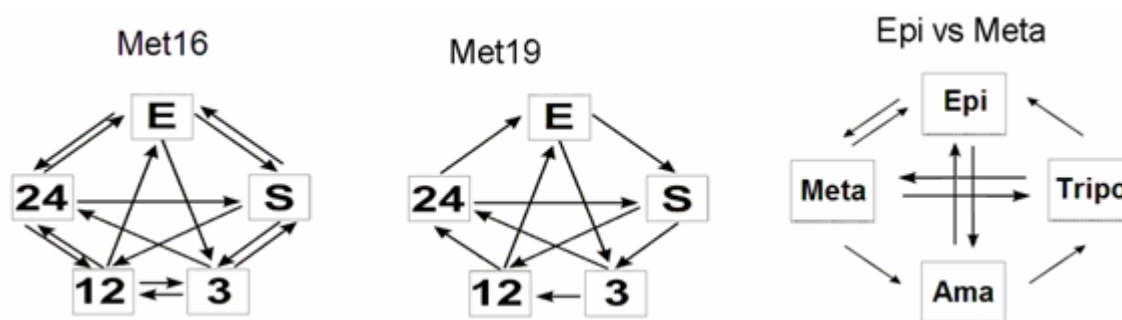


Figura 4.1 Desenho experimental das hibridações realizadas.
Met 16 e Met 19 são denominações de diferentes ensaios de metaciclôgênese

Cada uma das unidades demonstradas na Figura 4.1 foi feita para as amostras de mRNA total e polissomal correspondentes. A descrição de cada uma das unidades experimentais é a seguinte:

Met19: nessa réplica biológica da metaciclôgênese, foi utilizado o desenho em *loop* clássico, descrito por KERR & CHURCHILL (2001). Nesse desenho experimental, cada amostra é hibridada quatro vezes, com cada uma das outras quatro amostras do desenho, sendo que o número de marcações é balanceado, isto é, cada amostra foi marcada o mesmo número de vezes com cada um dos dois fluoróforos.

Met16: para essa réplica biológica da metaciclôgênese, foi acrescentado ao *loop* clássico uma hibridação em cada uma das amostras da parte externa do laço, criando uma réplica em *dye swap* de cada uma das amostras temporalmente adjacentes.

Epi versus Meta: esse mesmo desenho de *loop* foi repetido para as duas réplicas biológicas do experimento. A comparação entre Epi e Meta foi incluída dentro de um contexto maior de comparação das formas principais do ciclo de vida de *T. cruzi* (PROBST *et al.*, em preparação). Como a comparação das formas metacíclica versus epimastigotas era de interesse do presente projeto, foi incluída uma hibridação a mais, criando um *dye swap* entre essas amostras.

4.4 Fluxograma das etapas desenvolvidas

No intuito de aumentar o conhecimento sobre proteínas hipotéticas diferencialmente expressas durante a metaciclogênese, criou-se um conjunto de etapas laboratoriais sequenciais a serem seguidas, cujo fluxograma está representado na Figura 4.2.

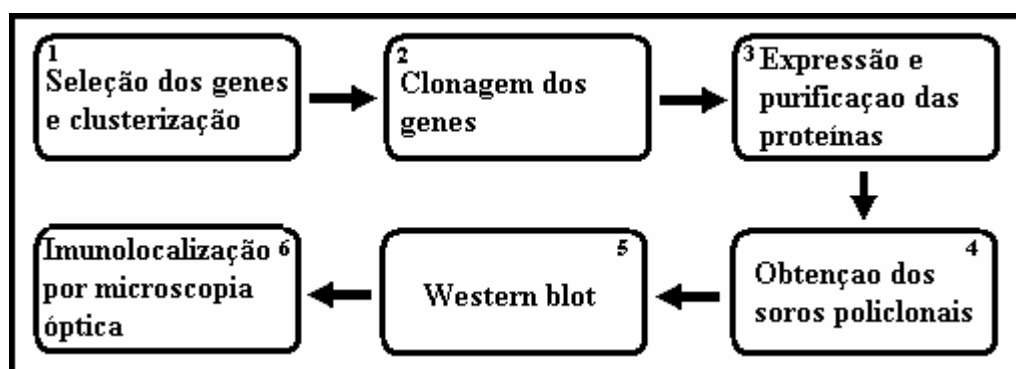


Figura 4.2. Fluxograma descritivo das etapas laboratoriais realizadas.

4.5 Análise dos dados gerados pelo microarranjo de DNA e seleção dos genes

As análises dos dados das hibridizações foram realizadas pelo setor de bioinformática do nosso instituto, e baseiam-se em uma lista contendo as sondas que possuíram um perfil diferencial de expressão, estando nelas contidos genes com uma expressão aumentada ou diminuída no mínimo duas vezes nas formas tripomastigotas metacíclicas em relação às formas epimastigotas.

Nesta lista estão contidos 912 sondas diferencialmente expressas conforme relatado anteriormente, das quais foram selecionados os genes avaliados no presente trabalho.

Inicialmente, foram selecionados sete genes representados na Tabela 4.1, denominados principais. A seleção destes genes foi realizada de acordo com o perfil de expressão durante a metaciclogênese e características da sonda, ou seja:

- a sonda deveria estar diferencialmente expressa em metacíclicos, com um grau de modulação de pelo menos duas vezes;
- o perfil da curva de hibridização durante a metaciclogênese deveria se apresentar de forma coerente entre os diferentes pontos, evitando-se erros técnicos dependentes da amostra;

- a proteína codificada deveria ser classificada como hipotética conservada, isto é, sua função não foi atribuída e deve estar presente em outros organismos;
- a proteína deveria possuir um módulo conservado indicativo de função;

Após a escolha dos genes principais, os dados relativos a todos os genes diferencialmente expressos dentro dos experimentos realizados no IBMP foram organizados pelo uso do programa de agrupamento Cluster 3.0 (DE HOON *et al.*, 2004), visando identificar grupos gênicos co-regulados, de acordo com os padrões gerais de expressão conforme exemplificado na Figura 4.3.

Dentro do grupo de cada gene principal, foram selecionados mais três genes distintos. Com essa seleção, foram incluídos 21 novos genes, denominados secundários. O perfil de cada cluster obtido pode ser verificados nas Figura 4.4 a Figura 4.10

Portanto, há sete conjuntos de quatro tipos distintos de genes, os quais podem ser caracterizados da seguinte forma:

Gene principal: codifica uma proteína hipotética conservada com módulo funcional identificado.

Gene secundário: representa uma redundância, também codificando uma proteína hipotética conservada com módulo funcional identificado. Constitui-se em uma proteção nos casos no qual os dados para o gene principal não forem obtidos e, quando não for esse o caso, representa um segundo elemento co-regulado cuja caracterização funcional é de interesse.

Gene guia: codifica uma proteína cuja caracterização funcional é mais conhecida, representando uma referência funcional para o cluster estudado. No entanto, não deve haver caracterização funcional desse gene dentro do processo de metaciclologênese, possibilitando que os resultados obtidos no presente trabalho sejam inéditos;

Gene acessório: codifica uma proteína hipotética, não classificada como conservada. Esse grupo contém as proteínas cuja relação de ortologia com outros organismos não foi identificada, podendo representar genes restritos a *T. cruzi* e, portanto, potencialmente de grande interesse para o estudo de questões biológicas peculiares a esse parasita. No entanto, também podem representar um erro de anotação, pela inexistência de corroboração em outros organismos, consistindo em uma proteína que não existe no organismo estudado.

Tabela 4.1. Descrição dos genes selecionados para caracterização e sua classificação interna

Grupo	ID IBMP	Gene ID	Classificação
1	TcExp02A01	8530.t00003	Gene Principal
	TcExp02A08	8694.t00004	Gene Acessório
	TcExp02A09	7626.t00009	Gene Secundário
	TcExp02A10	6680.t00001	Gene Guia
2	TcExp02A02	6996.t00063	Gene Principal
	TcExp02A11	8773.t00014	Gene Acessório
	TcExp02A12	8364.t00002	Gene Secundário
	TcExp02B01	8208.t00006	Gene Guia
3	TcExp02A03	7109.t00005	Gene Principal
	TcExp02B02	8764.t00016	Gene Acessório
	TcExp02B03	8728.t00021	Gene Secundário
	TcExp02B04	8257.t00019	Gene Guia
4	TcExp02A04	6896.t00027	Gene Principal
	TcExp02B05	7224.t00006	Gene Acessório
	TcExp02B06	6958.t00014	Gene Secundário
	TcExp02B07	8243.t00004	Gene Guia
5	TcExp02A05	4715.t00001	Gene Principal
	TcExp02B08	5595.t00003	Gene Acessório
	TcExp02B09	8171.t00006	Gene Secundário
	TcExp02B10	8416.t00006	Gene Guia
6	TcExp02A06	8415.t00003	Gene Principal
	TcExp02B11	7000.t00002	Gene Acessório
	TcExp02B12	8318.t00003	Gene Secundário
	TcExp02C01	7168.t00007	Gene Guia
7	TcExp02A07	5323.t00004	Gene Principal
	TcExp02C02	8359.t00034	Gene Acessório
	TcExp02C03	7741.t00008	Gene Secundário
	TcExp02C04	6996.t00038	Gene Guia

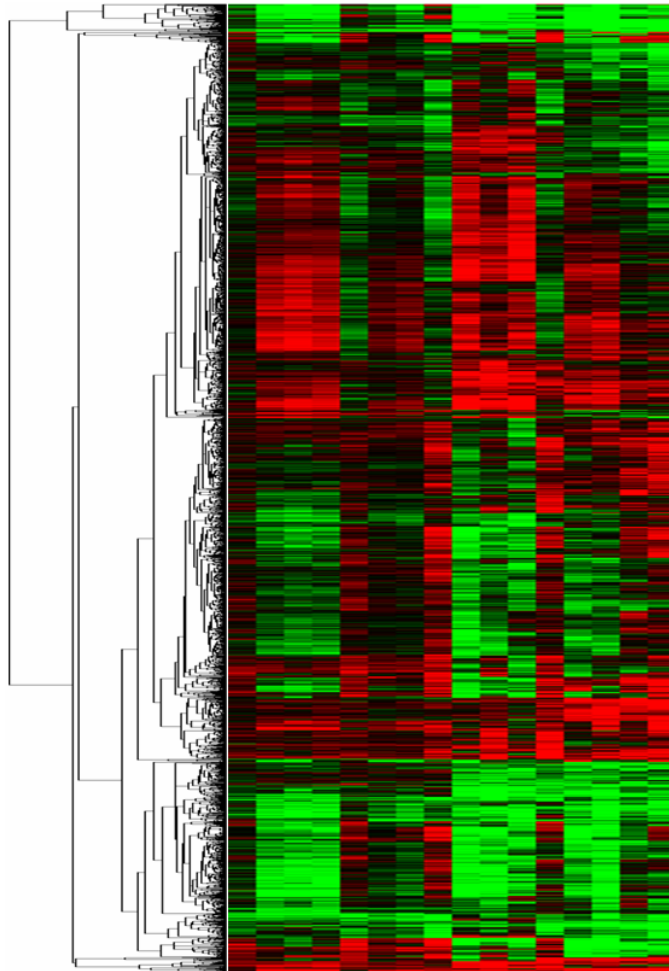


Figura 4.3 Agrupamento hierárquico dos dados de microarranjo disponíveis no IBMP.

À esquerda, dendrograma representando a similaridade dos genes (linhas); cada coluna representa um ponto experimental distinto. Em preto, expressão igual, em vermelho, expressão aumentada; em verde, expressão diminuída.

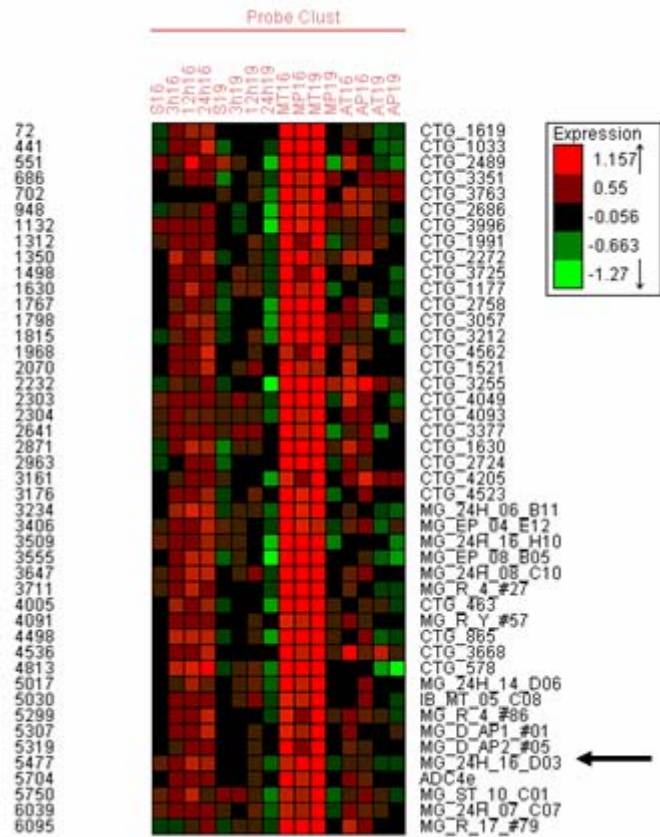


Figura 4.4 Cluster contendo a sonda referente ao gene 4715.t00001, grupo 5.
Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene seleccionado.

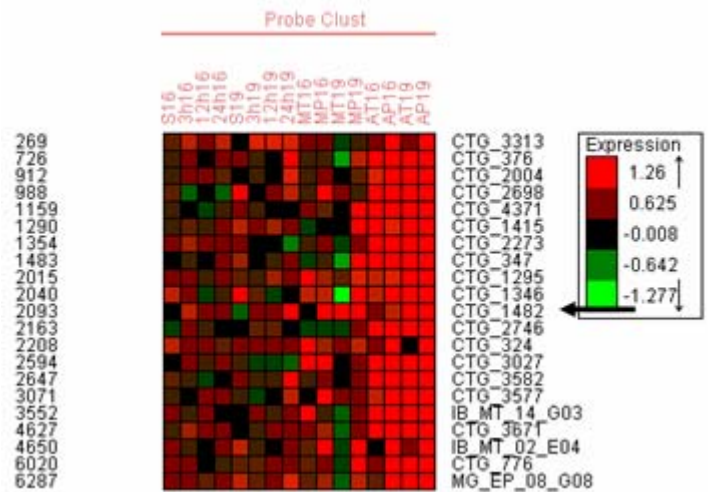


Figura 4.5 Cluster contendo a sonda referente ao gene 5323.t00004, grupo 7.
Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene seleccionado.

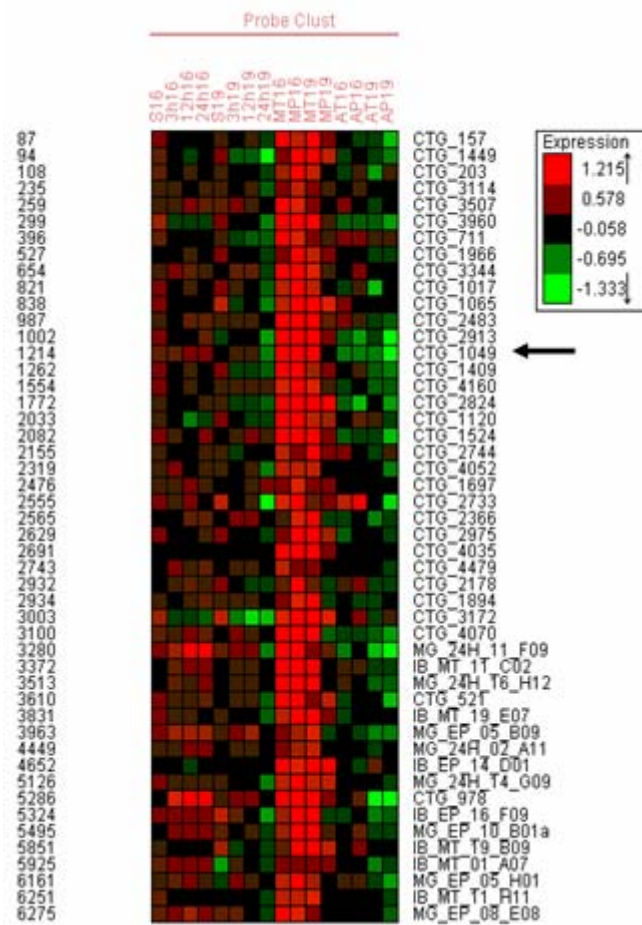


Figura 4.6 Cluster contendo a sonda referente ao gene 6896.t00027, grupo 4. Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene selecionado.

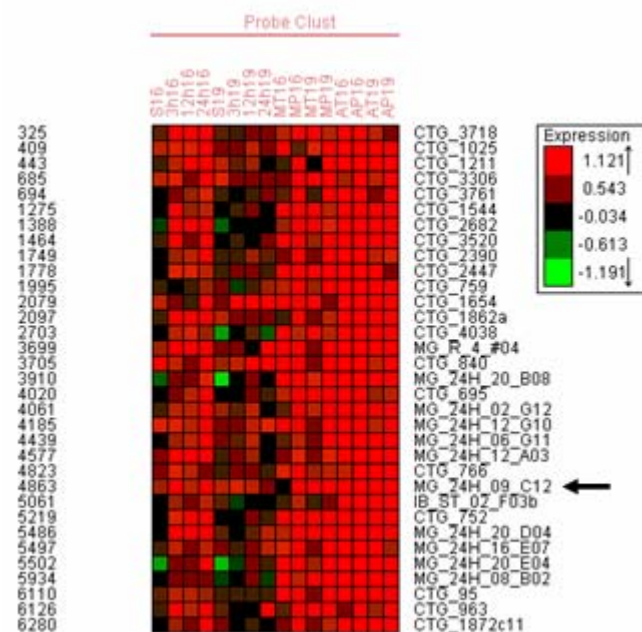


Figura 4.7 Cluster contendo a sonda referente ao gene 6996.t00063, grupo 2. Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene selecionado.

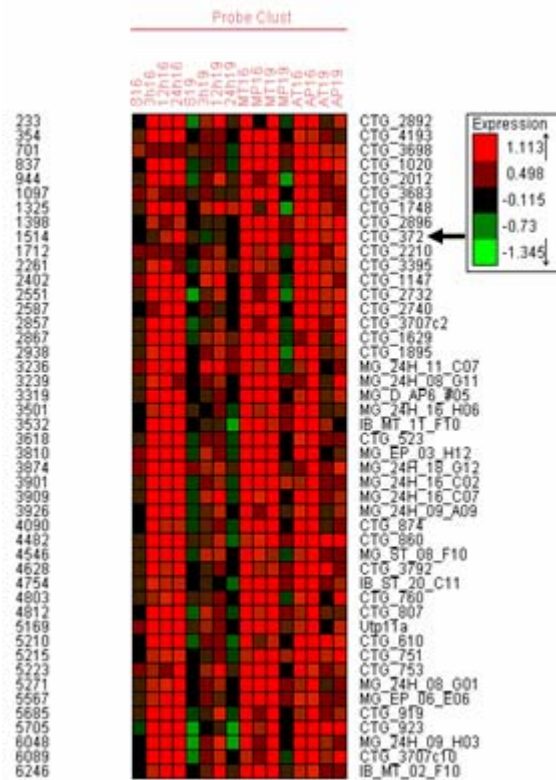


Figura 4.8 Cluster contendo a sonda referente ao gene 7109.t00005, grupo 3.
Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene selecionado

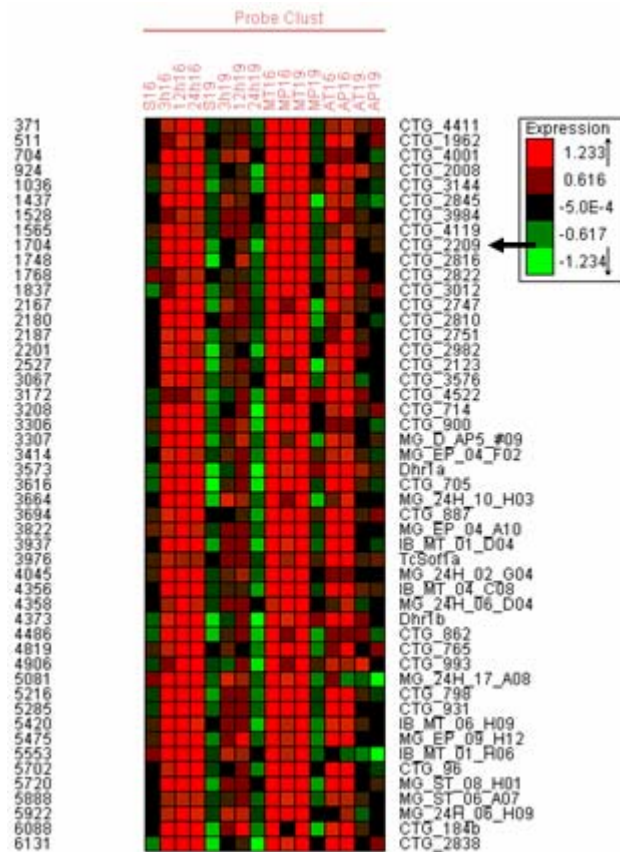


Figura 4.9 Cluster contendo a sonda referente ao gene 8415.t00003, grupo 6.
Descrição da visualização veja Figura 4.3. Seta indica sonda relativa ao gene selecionado

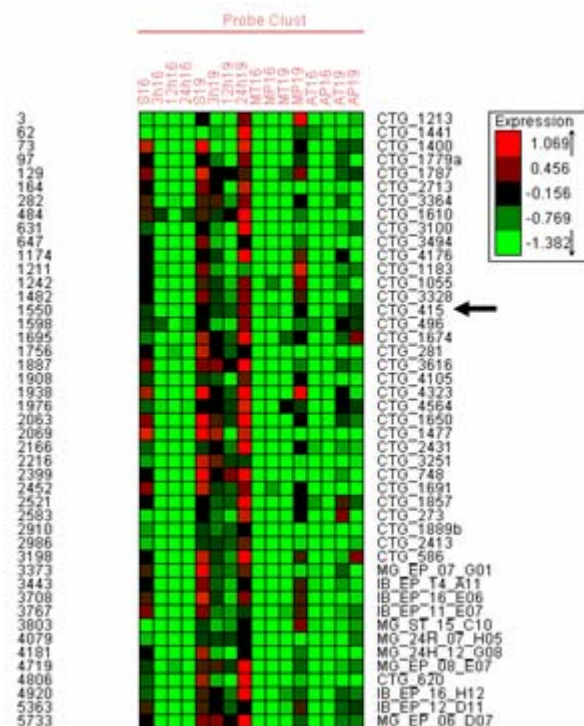


Figura 4.10 Cluster contendo a sonda referente ao gene 8530.t00003, grupo 1.
Descrição da visualização veja Figura 4.3. Sta indica sonda relativa ao gene selecionado

Baseado nos critérios acima citados, os genes foram selecionados e foram desenhados iniciadores para amplificação de toda a região codificadora, através do uso do software Primer Select (Lasergene Inc.), com a inclusão de adaptadores específicos da plataforma Gateway[®], conforme detalhado na Tabela 4.2.

Tabela 4.2: Iniciadores construídos adaptados para plataforma Gateway®.

INICIADORES	TM	GENE	Nome Iniciador
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCCTGCGTATCCGTC	57.0	8530.t00003	TcExp02A01F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCACTACACGGCGATCATTGCG	57.3	8530.t00003	TcExp02A01R1
GGGGACCACTTTGTACAAGAAAGCTGGGTCAACGCCACGGGAGCCAC	58.8	8530.t00003	TcExp02A01R2
GGGGACCACTTTGTACAAGAAAGCTGGGTCAAGCAGCATCATTACGGGAAC	51.7	8530.t00004	TcExp02A01R3
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCCGATGCCTCTATTTCGTATTG	55.3	6996.t00063	TcExp02A02F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGTGTTTTGAGTAACTCTGTGGATTCT	54.9	6996.t00063	TcExp02A02R1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTCAGTGACAAAACACACATGCG	54.4	7109.t00005	TcExp02A03F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTATTAGCCATCAGAGGAATC	53.1	7109.t00005	TcExp02A03R1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTCCCGAAGCGTGCCTTG	55.3	6896.t00027	TcExp02A04F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGGCGCCTCCGTAACAAAATC	56.0	6896.t00027	TcExp02A04R1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTCGAACGAGAGTTGAAGTGGC	53.4	4715.t00001	TcExp02A05F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCCGTTTAAACCCTCTTGAGTGGC	54.6	4715.t00001	TcExp02A05R1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTATGTGGCGTTTCTTGATGGC	54.8	8415.t00003	TcExp02A06F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCAAGACTTCTGGTGTACTTGGACTTCAC	55.4	8415.t00003	TcExp02A06R2
GGGGACCACTTTGTACAAGAAAGCTGGGTCAAGACTTCTGGTGTACTTGGCCTTC	56.1	8415.t00003	TcExp02A06R1
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTCATCTGAAATAGAACCGCATA	49.5	5323.t00004	TcExp02A07F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTTGCATGGTTCGTGAA	49.3	5323.t00004	TcExp02A07R1

Tabela 4.2 (continuação)

GGGGACAAGTTTGTACAAAAAGCAGGCTTCTGTTTTATGCGTGGAGAAG	52.4	8694.t00004	TcExp02A08F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGACTTGCCGCCAATAAGC	50.1	8694.t00004	TcExp02A08R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCCAATGTCAAGGGAATTCTCACTG	54.5	7626.t00009	TcExp02A09F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGCGCGGGGTGTCGC	55	7626.t00009	TcExp02A09R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGTCAGCTGGGGTCTCTG	53.9	6680.t00001	TcExp02A10F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTCTCTCGGCTTTTACCAG	54.8	6680.t00001	TcExp02A10R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCGCGCAGAGTCCGACTCTTATC	57.7	8773.t00014	TcExp02A11F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGCACTTCTATTGCGCACAATTG	58.4	8773.t00014	TcExp02A11R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCTCTTTAGTTACGATAATGTGGTTGC	53.2	8364.t00002	TcExp02A12F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTACCGTGGCGGGGAGC	54.6	8364.t00002	TcExp02A12R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATAGCGTCGGCAACGCCA	57.4	8208.t00006	TcExp02B01F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTACAAGAGCATGTACCGACGTCT	56.6	8208.t00006	TcExp02B01R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATCGCATCCGCTACGCCACACGCA	57.4	8208.t00006	TcExp02B01FMut
GGGGACCACTTTGTACAAGAAAGCTGGGTCTACAGGAGCATGTACCATTGTCTTTGAAGG	56.5	8208.t00006	TcExp02B01RMut
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGAGTGTGAGCACCAGGACAAGCG	55.9	8764.t00016	TcExp02B02F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGAGTCTTTTCTCCCTCCACCAG	55	8764.t00016	TcExp02B02R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCGACGAATGAGTGGTACGACATGC	55.9	8728.t00021	TcExp02B03F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGACCTCTTCAATTTTGTCCGAG	56.7	8728.t00021	TcExp02B03R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGCACCAACATCATGGATG	58.9	8257.t00019	TcExp02B04F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGCATCTATTCTTCTTTGTGCTCG	58.6	8257.t00019	TcExp02B04R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCGACGAGACTTCGCTCCCATC	53.6	7224.t00006	TcExp02B05F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGAGTCCGTGCCGTTGTG	53.3	7224.t00006	TcExp02B05R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGACAGTTCCTACTACGAGGAG	56.2	6958.t00014	TcExp02B06F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGACGGAGACAAGGCAGAGTTCAC	56.7	7224.t00006	TcExp02B06R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGAATACATCGGGCTCAGG	57.7	8243.t00004	TcExp02B07F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTGTGTTGATCTCTGTACTACAACCAATAC	56.5	8243.t00004	TcExp02B07R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCGAAATACATCGGGCTCAGG	50.6	8243.t00004	TcExp02B07F1Mut
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTGTGTTGATCTCTGTACTACAACCAATAC	51.1	8243.t00004	TcExp02B07R1Mut
GGGGACAAGTTTGTACAAAAAGCAGGCTTCGACTTTTTACTGTGCTGAACCAAC	56	5595.t00003	TcExp02B08F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCAAGTAAGGGGCTAATCAAACGACTC	55.4	5595.t00003	TcExp02B08R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGACGGGAGTGTGTTGGG	54.4	8171.t00006	TcExp02B09F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGAAGAATTTATAAATGAGCGAAGG	54.9	8171.t00006	TcExp02B09R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGGCGAGGGGTCTTCC	55.9	8416.t00006	TcExp02B10F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCAATGGAGAGGCTAAGGCGATCTC	56.9	8416.t00006	TcExp02B10R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCCGTTCTTGTCTAGAGGCCGAATG	58.4	7000.t00002	TcExp02B11F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGGTAAGAAAACCACTTCTAGAGGGTAC	57.4	7000.t00002	TcExp02B11R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCCGATCTTGCCAGCGACCGAATGTG	55.7	7000.t00002	TcExp02B11F2Mut
GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGGTAAGAAAACCACTTCCAGAGGG	54.2	7000.t00002	TcExp02B11R2Mut
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGCCGGAGGAGAAGGACC	55.2	8318.t00003	TcExp02B12F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTATGGCAGCAGTAGTGACGAG	54.5	8318.t00003	TcExp02B12R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGCGTCACGGGATCGTAC	56.9	7168.t00007	TcExp02C01F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGTCAACGGCTTCAAAAAGAG	56.3	7168.t00007	TcExp02C01R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGACGCGACGTCCGAACTC	57.3	8359.t00034	TcExp02C02F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTATTACAGATGTGCCAGGCACGC	58.6	8359.t00034	TcExp02C02R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGACGCGACGACCGAACTCTG	53.4	8359.t00034	TcExp02C02F1Mut
GGGGACCACTTTGTACAAGAAAGCTGGGTCTATTACAAATATGCCACGCTCTCTTATGTCC	53.8	8359.t00034	TcExp02C02R1Mut
GGGGACAAGTTTGTACAAAAAGCAGGCTTCCAGGATCCCATGGAGTTTATGAG	54.2	7741.t00008	TcExp02C03F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGGCAGTATACGTAACGGGCATC	55.2	7741.t00008	TcExp02C03R1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGGCAGTATACGTAACGCAATTCACAC	54.8	7741.t00008	TcExp02C03MutR1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCTCGACAACACGCGTGCTG	59.1	6996.t00038	TcExp02C04F1
GGGGACCACTTTGTACAAGAAAGCTGGGTCTGTAATCCTCATCCTCTAGCATGATGGC	59.4	6996.t00038	TcExp02C04R1
GGGGACAAGTTTGTACAAAAAGCAGGCTTCTCGACAACAAGAGCAGCGGG	53.4	6996.t00038	TcExp02C04F1Mut
GGGGACCACTTTGTACAAGAAAGCTGGGTCTAATCTTCGTCTCTAACATGATGGCAAG	52.4	6996.t00038	TcExp02C04R1Mut

Iniciadores das seqüências nucleotídicas com adaptadores da plataforma Gateway (em negrito está a seqüência específica ao gene); TM (temperature melting); Gene (identificação do gene); Nome (controle interno).

4.6 Amplificação dos genes

Para a amplificação dos genes, foram realizados testes objetivando a melhor performance da reação, como utilização de gradiente de temperatura e concentração de reagentes. Neste processo de testes e também na amplificação dos genes, foi utilizada a enzima *High-Fidelity Triple Master Enzyme* (Eppendorf). A partir da padronização das condições de PCR, deu-se início à amplificação dos genes.

As condições de amplificação foram as seguintes:

- 94°C por dois minutos;
- 10 ciclos de 94°C (30 segundos), 57°C (30 segundos) e 72°C (90 segundos);
- 25 ciclos de 94°C (30 segundos), 62°C (30 segundos) e 72°C (90 segundos).

Para os genes com tamanho maior de 1.500 pares de bases, o tempo de extensão foi aumentado para 3 minutos.

As concentrações dos reagentes utilizados foram:

- Iniciadores *forward* e *reverse* (5 μ M);
- DNA genômico (100 ng);
- 10x Tampão (*High Fidelity Buffer*) com Mg^{++} (2,5 mM);
- dNTP (200 μ M);
- Enzima *Triple Master* (0,05 U/ μ l);
- Água q.s.p. 50 μ l.

Para os genes com tamanho maior de 1.500 pares de bases, a concentração da enzima *Triple Máster* foi aumentada para 0,1 U/ μ l.

Após a amplificação dos produtos foram realizadas eletroforeses em gel de agarose para verificação do qualidade da amplificação.

4.7 Purificação dos produtos de PCR

Para retirar pequenos fragmentos de DNA sintetizados e os iniciadores empregados, provenientes do processo de PCR, foi utilizada a purificação baseada em polietilenoglicol ou PEG (30% PEG 8000, 30mM $MgCl_2$). Ao produto de PCR acrescentou-se quatro volumes de tampão TE (10 mM Tris-HCl, 1 mM EDTA, pH 8,0). Em seguida acrescentou-se o PEG, 50%

v/v. Após, as soluções foram completamente homogêneas e colocadas sob centrifugação a 10.000 g por 15 minutos em temperatura ambiente. Após centrifugação, o sobrenadante foi retirado cuidadosamente e descartado, e o pellet ressuspensão em 10 µl de tampão TE. Após este processo de purificação, as amostras foram submetidas à eletroforese em gel de agarose e submetidas à dosagem utilizando espectrofotômetro.

4.8 Clonagem e vetores

Os produtos dos genes selecionados, após purificação, foram clonados em vetores do sistema de clonagem comercial Gateway® (Invitrogen), cuja tecnologia é baseada em recombinação utilizando seqüências sítio específicas do bacteriófago lambda (LANDY, 1989), possibilitando uma rápida e eficiente análise funcional das proteínas de interesse (HARTLEY *et al*,2000).

O sistema Gateway® permite que, a partir do clone de entrada, seja feita a passagem do inserto, de forma rápida e eficiente, para um segundo vetor, denominado de destino, que pode ser dos mais variados tipos, de acordo com a necessidade de caracterização funcional, como mostra a Figura 4.11.

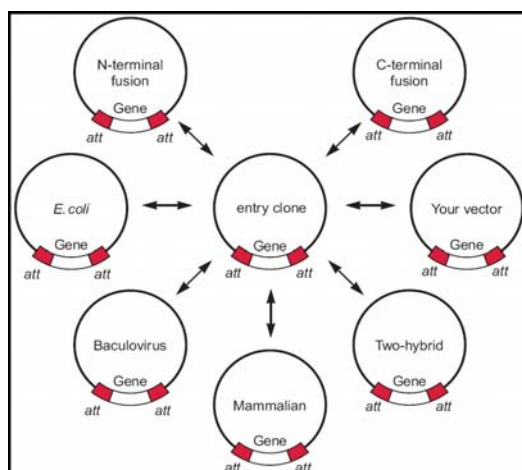


Figura 4.11 Possibilidades de utilização do sistema Gateway a partir da obtenção do clone de entrada.

(Fonte: www.invitrogen.com)

A recombinação entre os sítios do bacteriófago lambda acontece tanto no ciclo lisogênico, entre os sítios attB e attP, resultando nos sítios attL e attR, quanto no ciclo lítico, entre os sítios attL e attR, resultando novamente nos sítios attB e attP. Essas reações são específicas e direcionais, pois attB1 e attP1, attB2 e attP2, attL1 e attR1, attL2 e attR2

recombinam somente entre si.(Figura 4.12 e Figura 4.13). A enzima utilizada na reação BP na verdade é formada por uma integrase do bacteriófago lambda (Int) e um fator de integração ao hospedeiro de *E. coli* (IHF).

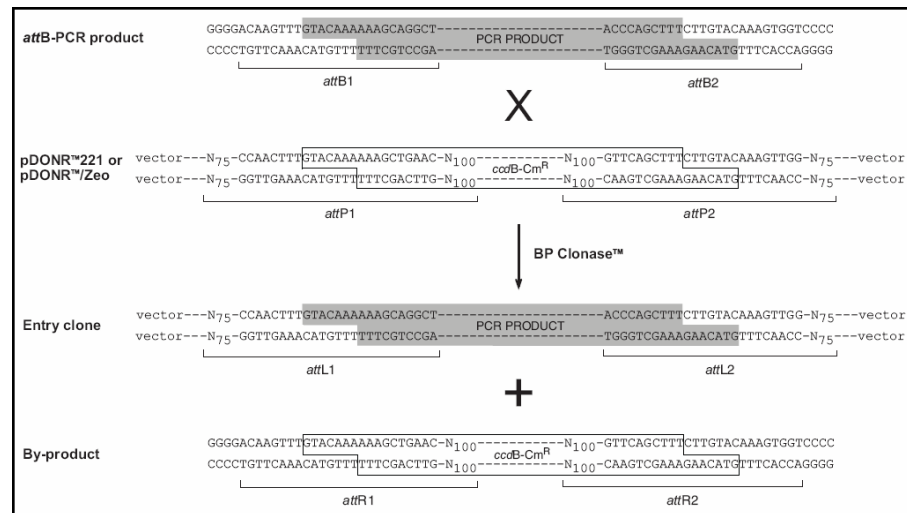


Figura 4.12 Reação attB x attP.

O produto específico de pcr contendo o sítio attB1 e 2 é inserido no vetor escolhido (pDONR™221) através da utilização da enzima BP clonase, recombinao com o sítio attP1 e 2 presente no vetor, formando um novo sítio denominado de attL 1 e 2 no clone selecionado por antibiótico.

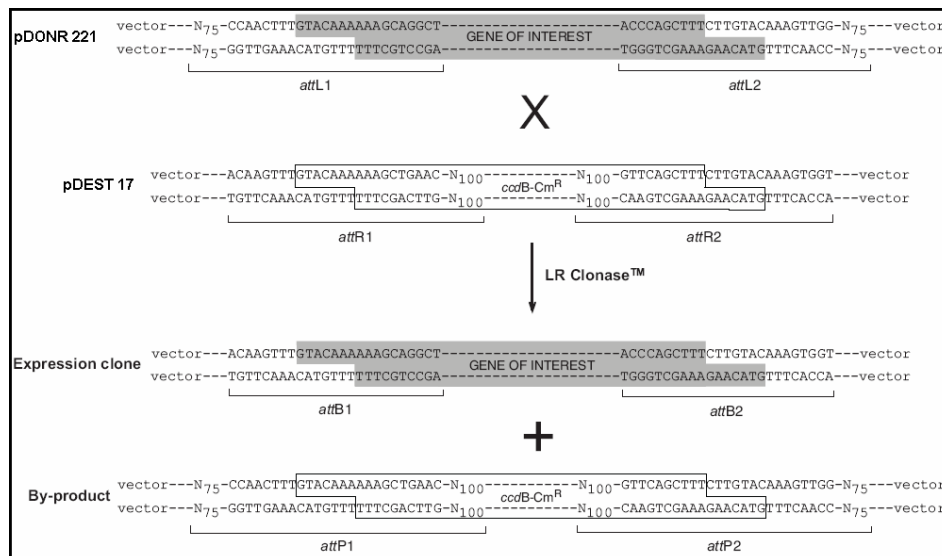


Figura 4.13 Reação attL x attR

O vetor de entrada contendo a sequência do gene de interesse o qual possui o sítio attL 1 e 2 é recombinao com o vetor de destinao (pDEST™17) através da utilização da enzima LR clonase, recombinao com o sítio attR1 e 2 presente no mesmo.esta recombinao forma novamente o sítio attB 1 e 2 no clone de destinao selecionado por antibiótico.

Para que um produto amplificado a partir de um determinado gene de interesse ingresse nesta plataforma, escolhemos o vetor de entrada pDONR™221 (Invitrogen, Figura 4.14) que possibilita a clonagem direcional de produtos de PCR.

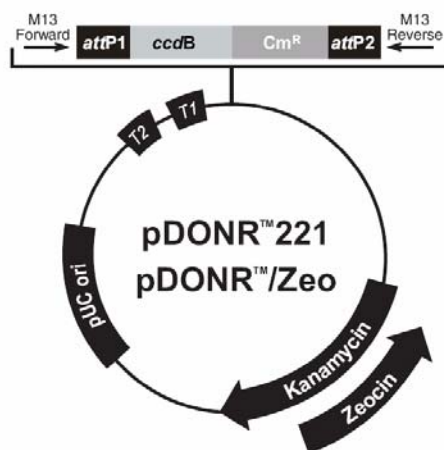


Figura 4.14 Vetor de entrada utilizado, pDONR 221(Invitrogen).

A partir da obtenção dos clones de entrada, é possível inseri-los em diversas plataformas. No presente trabalho, foi realizado a recombinação com o vetor pDESTTM17 (Figura 4.15), o qual é apropriado para a expressão heteróloga da proteína de interesse em *E. coli*, fusionada a uma sequência de seis histidinas em sua extremidade N-terminal.

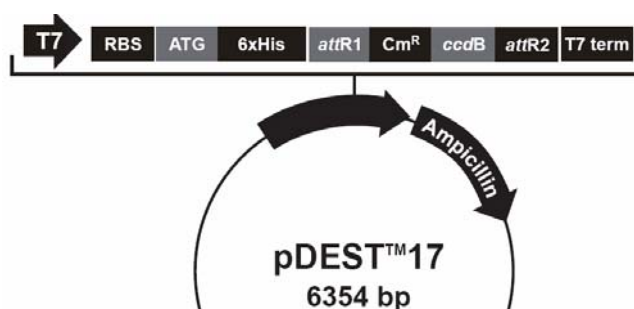


Figura 4.15 Vetor de expressão pDESTTM17 (Invitrogen).

4.8.1 Obtenção dos clones de entrada

A inserção dos produtos de PCR no vetor de entrada se baseia na recombinação dos sítios conforme descrito anteriormente. Aos produtos foram acrescentadas as sequências que possuem os sítios de recombinação durante o processo de PCR, os quais recombinarão com os sítios presentes no vetor pDONRTM221.

A reação foi realizada conforme orientações do fabricante (*Gateway BP clonase 2 Enzyme mix*, Invitrogen). O processo consistiu em acrescentar 150 ng do produto de PCR contendo o sítio attB, 150 ng do plasmídeo pDONRTM221, e completando até 8 microlitros com TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8). As reações foram incubadas a 25°C

por duas horas. Após esta etapa, foi adicionado 2 µg de proteinase K e a reação incubada por 10 minutos a 37°C.

Após a conclusão dessa etapa, foi realizada a transformação em células cálcio-competente da linhagem de *E.coli* DH5α. Acrescentou-se 1 µl da reação BP a 50 µl de células cálcio-competentes. Estas células permaneceram no gelo por 30 minutos e, em seguida, foram incubadas a 42° C por 2 minutos, sendo colocadas novamente no gelo.

Após a transformação, as células cresceram em 1 ml de meio LB por uma hora a 37 °C, para em seguida serem centrifugadas a 4000 g por 5 minutos e colocadas para crescer em placa de LB contendo kanamicina (25mg/ml) a 37° C durante 15 horas.

Para verificar a presença de clones com os insertos de interesse, foram feitas *master plates* e PCR de colônia, para a qual o protocolo seguido foi conforme o fabricante da enzima Taq polimerase (Invitrogen). Após a finalização da PCR, foi realizada eletroforese em gel de agarose para verificação e escolhas das colônias contendo os clones com o inserto correto. As colônias selecionadas foram colocadas em placa contendo LB e kanamicina (25 mg/ml), por 15 horas a 37° C. Uma colônia de cada placa referente a cada um dos genes foi selecionada e colocada em 10 ml de meio LB contendo kanamicina (25 mg/ml), por 15 horas a 37° C sob agitação (220 RPM).

Após o crescimento em cultura do clone bacteriano, foi feita a purificação do plasmídeo utilizando-se o produto comercial *QIAprep Spin Miniprep* (Quiagem) conforme orientações do fabricante. Após a obtenção dos plasmídeos, fez-se a mensuração da concentração através do uso de espectrofotômetro.

4.8.2 Obtenção dos clones de expressão

A troca de insertos entre os plasmídeos de entrada e destino, pDONR e pDEST, denominada de reação LR, foi realizada conforme orientações do fabricante (Gateway LR clonase 2 Enzyme mix, Invitrogen). Foi utilizado 150 ng de pDONR contendo o gene de interesse, 150 ng do plasmídeo pDEST, e completando o volume até oito microlitros com TE (10 mM Tris-HCl, 1 mM EDTA, pH 8). A reação foi incubada a 25°C por duas horas.. Após esta etapa, foi adicionado 2 µg de proteinase K e a reação incubada por 10 minutos a 37°C.

Após essa etapa, foi realizada a transformação de células cálcio-competentes da linhagem de *E.coli* DH5α, conforme descrito anteriormente.

Após a transformação as células foram colocadas em 1 ml de meio LB por uma hora a 37 °C, para em seguida serem centrifugadas a 4000 g por 5 minutos e colocadas para crescer em placa de LB contendo ampicilina (100 mg/l) a 37° C durante 15 horas.

A verificação da obtenção dos clones de destino foi realizada conforme descrito anteriormente. As colônias selecionadas foram estriadas em placa contendo LB e ampicilina (100 mg/l), a qual foi colocada em estufa a 37° C durante 15 horas. Uma colônia de cada placa referente a cada um dos genes foi selecionada e colocada em 10 ml de meio LB contendo ampicilina (100 mg/l), para obtenção dos plasmídeos, e deixar crescendo por 15 horas a 37° C sob agitação (220 RPM).

A purificação do plasmídeo foi realizada conforme descrito anteriormente.

4.9 Expressão das proteínas

Esta etapa visou a obtenção da proteína visando a inoculação em camundongos, para obtenção de anticorpos policlonais. As células utilizadas para esta etapa foram da linhagem de *E.coli* BL21(DE3) pLysE. A expressão das proteínas teve início com a transformação por choque térmico das células. Em 50 µl de células, foram adicionados 300 ng de plasmídeo de expressão contendo o gene de interesse. As células foram então mantidas por 30 minutos no gelo e em seguida incubadas a 42° C por dois minutos voltando ao gelo por mais três minutos. Após isto, foi acrescentado ao tubo 1 ml de meio LB, com posterior incubação a 37° C por 1 hora, sob agitação de 220 RPM. Após, as células foram inoculadas em meio LB contendo cloranfenicol (25mg/l) e ampicilina (100mg/l) em uma diluição de 1/10 e deixado sob agitação de 220 RPM a 37° C, durante 15 horas. Após este período, foi feito um novo inóculo em uma diluição de 1/10 em um volume final de 250, 500 ou 1000 ml.

O inóculo foi colocado em meio LB contendo cloranfenicol (25mg/l) e ampicilina (100 mg/l). Após uma hora e meia a densidade óptica começou a ser monitorada, até que se obtivesse o valor de 0,8 a 600 nm, sendo então adicionado IPTG (0,1 mM) visando a indução da expressão da proteínas desejada. Imediatamente antes desta etapa, foi retirada uma alíquota de 1 ml, a qual foi denominada de amostra não induzida. Esta amostra, juntamente com a cultura induzida, foi colocada sob agitação a 37° C por duas horas.

Após duas horas de indução, a cultura foi centrifugada a 5000 RPM durante 15 minutos a 4° C; o sobrenadante foi descartado e o *pellet* ressuspenso em tampão de sonicação (Tris-HCl 50 mM pH 8). Feito a ressuspensão, as células foram sonicadas por 20 segundos (potência 8),

repetindo-se esse processo por mais quatro vezes. Após esse procedimento, foi retirada uma alíquota de 50 µl, denominada de amostra total. O restante foi centrifugado a 10.000 g por 15 minutos a 4°C para obtenção das frações denominadas solúveis e insolúveis.

A fração solúvel é constituída pelo sobrenadante obtido após a centrifugação, o qual foi armazenado a -20°C. O *pellet* resultante foi ressuspenso em tampão de solubilização (uréia 8 M, imidazol 10 mM, NaCl 300 mM) e em seguida sonicado, conforme descrito acima, e submetido a centrifugação (10.000 g por 15 minutos a 4°C) sendo então, o sobrenadante denominado fração insolúvel, e armazenado a -20°C.

4.10 Comprovação da expressão

A verificação da expressão foi realizada através da técnica de SDS-PAGE com diferentes concentrações de acordo com o peso molecular das proteínas, utilizando 30 mA de corrente elétrica por gel. Os géis foram corados com azul de comassie por 30 minutos, sendo retirado o excesso com solução de descoloração (metanol 30 %, ácido acético 10 %). Para a amostra controle, a bactéria BL21(DE3) pLysE cresceu sem transformação em meio LB. As células foram centrifugadas a 4.000 g durante 15 minutos a 4 °C e em seguida o *pellet* foi ressuspenso em 1 ml de tampão Tris-HCl 50mM pH 8 e sonicado 5 vezes por 20 segundos. Foi separado 100 µl e adicionado 100 µl de Tris-HCl 50mM pH 8 e ainda 0,5 µl de DNase NQI. O lisado resultante foi incubado no gelo por uma hora e em seguida adicionou-se 50 µl de tampão de desnaturação de proteína, e colocado em banho-maria a 90° durante dez minutos.

4.11 Purificação das proteínas

Após a obtenção das proteínas recombinantes e verificação da solubilidade, procedeu-se à purificação através de cromatografia de afinidade por níquel (QIAexpress System, Quiagem) e por eletroforese (SDS-PAGE).

O processo de purificação pode ser baseado em dois protocolos distintos de purificação, podendo ser feita eluição por imidazol ou por acidificação do pH

O processo de purificação baseado no imidazol inicia-se com a lavagem das colunas e equilíbrio das mesmas com a solução de lise (50 mM NaH₂PO₄, 300 mM NaCl e 10mM imidazol). Depois, os extratos protéicos são adicionados à coluna permitindo a fixação das proteínas contendo cauda de histidina à resina e a solução que não se liga é coletada e

armazenada, sendo denominada de *flow-through* (FT). Para retirar proteínas contaminantes da resina, foram feitas cinco lavagens com 10 ml de solução de lavagem (50 mM NaH_2PO_4 , 300 mM NaCl e 20mM imidazol), sendo que os materiais proveniente das lavagens foram também armazenados para verificação da qualidade do processo de purificação. Seguindo o processo de purificação foi feito então a eluição, sendo realizadas cinco eluições utilizando 0,5 ml de solução de eluição (50 mM NaH_2PO_4 , 300 mM NaCl e 250mM imidazol) e armazenadas a -20°C .

A purificação utilizando o decréscimo de pH inicia-se também com a lavagem das colunas e equilíbrio das mesmas com a solução de lise (100 mM NaH_2PO_4 , 10 mM Tris-HCl, 8 M uréia, pH 8). Depois os extratos protéicos são adicionados à coluna permitindo a fixação das proteínas, sendo o restante coletado e armazenado, sendo denominado de *flow-through* (FT). As lavagens foram feitas com 10 ml de solução de lavagem (100 mM NaH_2PO_4 , 10 mM Tris-HCl, 8 M uréia, pH 6,3), sendo que os materiais proveniente das lavagens armazenados para verificação da qualidade do processo de purificação e armazenados a -20°C . Para as eluições foi utilizado a solução de eluição (100 mM NaH_2PO_4 , 10 mM Tris-HCl, 8 M uréia) sendo que, a cada duas eluições com 0,5 ml da solução citada, o pH foi sendo reduzido, na seguinte ordem: 5,9; 4,5; 4,0 e 3,5. Os eluídos foram armazenados a -20°C .

Após cada processo de purificação, as amostras FT, os lavados e os eluídos foram analisados por SDS-PAGE para verificar a qualidade da purificação. A concentração de acrilamida utilizada variou entre 13 e 17% dependendo do tamanho esperado das proteínas, utilizando 30 mA de corrente elétrica por gel. Os géis foram corados com azul de comassie por 30 minutos, sendo retirado o excesso com solução de descoloração (metanol 30 %, ácido acético 10 %).

Após a realização da primeira etapa de purificação por afinidade, a segunda etapa baseada em separação por tamanho. Para isso, foi realizado um SDS PAGE em gel preparativo. A concentração de acrilamida utilizada variou entre 13 e 17% dependendo do tamanho esperado das proteínas, utilizando 15 mA de corrente elétrica por gel. Após o término, o gel foi submetido à coloração por KCl 1 M, e a banda desejada excisada do gel..

Para retirar a proteína desejada da matriz do gel de acrilamida foi utilizado o processo de eletroeluição. O material excisado foi colocado em uma membrana de diálise (Sigma-Aldrich) e adicionou-se 1 ml de tampão de SDS-PAGE (Tris 25mM, Glicina 192mM, SDS 0,1%) e lacrou-se as extremidades. Feito isto, as membranas foram colocadas dentro da cuba eletroforética contendo tampão SDS-PAGE e submetidas a uma corrente de 50 mA por 2

horas. Após este período, o líquido presente dentro da membrana foi coletado e o processo foi repetido.

A solução obtida foi submetida à eletroforese como descrito anteriormente. Com as imagens dos géis em mãos, foi realizada a dosagem com o uso do programa Lab Works 3.0.2 (UVP Inc). Este software permite a extração das intensidades de sinal das bandas obtidas permitindo assim a obtenção da concentração relativa de cada uma das proteínas, através da comparação com o marcador de peso molecular, do qual possui concentração conhecida (BenchMark Ladder™, Invitrogen).

4.12 Obtenção dos anticorpos policlonais.

As proteínas recombinantes foram inoculadas em camundongos da linhagem BALB/c, sendo que para cada proteína foram utilizados 2 camundongos, cada qual inoculado por uma diferente via: sub-cutânea ou intraperitoneal. Foram realizadas quatro inoculações de aproximadamente 50 µg de proteína em intervalos de 15 dias. A proteína foi emulsificada em adjuvante completo de Freund (Sigma) na primeira inoculação, e em halogel (Serva) nas inoculações subseqüentes. Sete dias após a última inoculação, procedeu-se à coleta do soro. Os animais foram sedados utilizando-se 0,2 mg de cetamina (Syntec) e 2 mg de xilasina (Vetbrands), coletando-se o sangue por punção cardíaca, o qual foi colocado por 5 minutos a 37°C e em seguida centrifugado por 10 minutos a 1.500 g, coletando-se o sobrenadante, o qual foi armazenado a -20° C.

4.13 Verificações da qualidade dos soros obtidos utilizando ensaios de Western blot.

Os soros obtidos foram testados para verificar a existência de anticorpos que reconheçam a proteína recombinante e também se possuem reação contra extrato protéico da fase epimastigota do *T. cruzi*.

A técnica de *Western blot* consiste na transferência de proteínas, previamente separadas por SDS-PAGE, para um suporte sólido (membrana de nitrocelulose). O procedimento a seguir foi executado segundo o método de TOWBIN *et al.* (1979). Após a eletroforese das proteínas recombinantes e do extrato de epimastigota de *T. cruzi*, as proteínas foram eletrotransferidas para uma membrana de nitrocelulose utilizando 50 mA à 4°C por 2 h. Após

a transferência, a membrana de nitrocelulose foi corada com solução de Ponceau S para verificar a qualidade da transferência sendo em seguida descorada com PBS/Tween 0,1% e incubado em solução de bloqueio (5% de leite desnatado em PBS/Tween 0,1%) por 16 horas a 4°C. Após essa incubação, a membrana foi colocada em 2 ml de solução de bloqueio contendo o anticorpo primário (policlonal obtido) diluído na proporção de 1:200 permanecendo por 1 hora a temperatura ambiente e, em seguida, lavada três vezes por 5 min com 50 ml de solução de bloqueio. Dois mililitros da solução contendo o anticorpo secundário (policlonal), anti-IgG de camundongo conjugado com fosfatase alcalina (Promega), diluído na proporção de 1:7.500, foram adicionados à membrana e incubou-se por 45 min. As membranas foram novamente lavadas 3 vezes (5 min) com 50 ml de PBS/Tween 0,1%. Para revelar as reações, foram utilizados 66 µl de NBT (cromógeno) e 33 µl de BCIP (substrato), diluídos em 10 ml de tampão para fosfatase alcalina.

4.14 Ensaios de Western blot para determinação de expressão

Após a confirmação da reatividade do soro com a proteína recombinante, foram feitas novas eletroforeses (SDS-PAGE) para cada um dos soros, utilizando extrato de *T. cruzi* das fases: epimastigota, fase intermediária da metaciclogênese (aderido 24 horas) e tripomastigota metacíclico. A partir dos géis obtidos na eletroforese, procedeu-se à realização da técnica de *Western blot* conforme descrito no item anterior.

Visando aumentar a quantidade de proteínas distintas avaliadas, a fim de possibilitar um conjunto de dados mais amplo, foram adicionados outros anti-soros disponíveis no IBMP, referentes às proteínas. PEPCCK (fosfoenolpiruvatocarboxiquinase, glicosomal, 7378.t00009/7730.t00002), Kap3 (proteína associada ao cinetoplasto 3, 6032.t00002/6142.t00008), Kap4 (proteína associada ao cinetoplasto 4, 5609.t00004/5609.t00005/5785.t00001/5785.t00003/5785.t00004), p22 (proteína associada ao kDNA p22, 5784.t00012/6036.t00001), proteína ribossomal L26 (8487.t00016/6039.t00002/8822.t00013), proteína ribossomal S7 (7013.t00007/7013.t00003/7108.t00012), Nop56 (proteína nucleolar, 4745.t00005/6845.t00001/6154.t00006), TcYchF (proteína de ligação ao RNA, 6380.t00002/8738.t00005), TcS1A12 (ciclofilina A, 7143.t00030), TcS1B11 (flavoproteína, 4913.t00016), TcS1B12 (isomerase de enoil-CoA, 5420.t00004/6028.t00004) e TcS2E02 (subunidade do proteassomo beta 6, 7369.t00005/7786.t00016).

A partir dos resultados obtidos, os soros que tiveram um padrão de reação fraco ou inexistente foram submetidos à técnica de coloração por quimioluminescência utilizando o produto ECL *Western blotting detection reagents and analysis system* (Amersham Biosciences), devido à sua maior sensibilidade em relação ao protocolo de detecção descrito no item 4.10. O protocolo foi seguido conforme orientações do fabricante.

4.15 Análise dos resultados obtidos a partir dos ensaios de Western blot.

A membrana revelada pelo método colorimétrico ou o filme revelado pelo método de quimioluminescência foram escaneados e a figura resultante foi analisada pelo programa Scion Image v. 4.0.3.2 (Scion Corporation).

As imagens foram pré-processadas, fazendo-se a subtração da intensidade de fundo (*background*) e a transformação da imagem, originalmente em escala de cores 24-bit para uma imagem binária (preto e branco), utilizando-se um limiar de intensidade que permitisse a correta identificação das bandas de interesse. Após essa etapa foram eliminados os pontos relativos a artefatos na imagem e procedeu-se à quantificação dos parâmetros: número de pixels da banda, intensidade mínima, máxima e média dos pixels, densidade integrada do sinal. Esse último parâmetro foi utilizado para as análises subsequentes de expressão protéica diferencial.

4.16 Imunolocalização por microscopia óptica

Foram realizadas imunofluorescências utilizando os anti-soros obtidos, usando a forma epimastigota do *T. cruzi*. Para cada anticorpo foi realizada uma diluição de 1:50 e 1:100. Alguns ensaios foram repetidos com uma diluição menor, de 1:50 e 1:25, a fim de melhorar a intensidade de sinal. Para os soros obtidos antes das inoculações das proteínas recombinantes (pré-imune) foi utilizada a menor diluição como controle.

Os parasitas foram centrifugados a 7.000 RPM por 3 minutos e ressuspensos em PBS, repetindo-se esse passo mais uma vez.

Acrescentou-se 50 µl de poli-L-lisina às lâminas e incubou-se a 37 °C por vinte minutos. Após este período, as lâminas foram lavadas com água ultra-pura e secas na estufa a 37°C. Os ensaios foram realizados em lâminas com dez campos limitados por teflon, sendo que cada

campo continha 2×10^6 parasitas, as quais foram colocadas a 28° C por uma hora em câmara úmida. Após este período de incubação o excesso foi retirado e adicionado 20µl de para-formaldeído por 15 minutos visando a fixação dos parasitas. Após, descartou-se o para-formaldeído e a lâmina foi lavada com PBS 1x; em seguida, acresceu-se 20µl de cloreto de amônio por 10 minutos para bloquear a fluorescência natural do para-formaldeído, e novamente procedeu-se à lavagem da lâmina em PBS 1x. A próxima etapa consistiu na adição de 20 µl de TRITON X-100 0,1% por 5 minutos e posterior lavagem por imersão, por 3 vezes, em PBS 1x, sendo em seguida, adicionados 20µl de PBS 1x contendo soro de cabra 25% por campo e deixadas por 16 horas a 4°C em câmara úmida, para realizar o bloqueio da lâmina. Após essa etapa, foi retirada a solução de bloqueio e acrescentado 20 µl de solução de bloqueio contendo o anticorpo primário (1/25, 1/50 e/ou 1/100), permanecendo em câmara úmida por uma hora a 37 °C. Após incubação, a lâmina foi lavada três vezes em PBS 1x por imersão por 5 min e acrescentado o anticorpo secundário (anti-camundongo, fluoróforo alexa) diluído (1:500) na solução de bloqueio, permanecendo por uma hora a 37 °C em câmara úmida, protegido da luz. Depois deste período, a lâmina foi lavada em PBS 1x por 5 minutos sendo repetida duas vezes o processo de lavagem, também protegido da luz, como todos passos a diante. Para a coloração de núcleo e cinetoplasto, foi acrescentado 20 µl de solução de bloqueio contendo DAPI em uma diluição 1:2000 permanecendo em temperatura ambiente por 5 minutos, sendo este processo, seguido de 10 lavagens por cinco minutos cada utilizando PBS 1X. Realizadas as lavagens foi adicionado a cada campo, 10 µl de n-propil-galato e as lâminas lacradas utilizando lamínula de microscopia e selador (esmalte).

As lâminas foram observadas no microscópio de fluorescência (Nikon). As imagens foram capturadas usando a câmara CoolSnap (Media Cybernetics) e analisadas com o programa Image Pro-Plus v. 4.5.1.22 (Media Cybernetics).

Utilizando o programa citado para análise, as imagens foram trabalhadas controlando brilho e contraste a fim de ressaltar os resultados para sua exibição. Além disso, as imagens relativas à marcação por DAP normalmente resultando em fluorescência azul, foram transformadas por alteração de canal para a cor vermelha, permitindo uma melhor sobreposição de imagens, devido ao excesso de background obtido nas imagens. Assim sendo, nas imagens mostradas no capítulo 5, a coloração vermelha nas fotos de imunofluorescência, é referente a marcação com DAPI, ou seja, marcação de DNA (cinetoplasto e núcleo).

5 Resultados

5.1 Descrição inicial

A caracterização do processo de metaciclogênese por microarranjo de DNA representa somente o passo inicial na elucidação de seus mecanismos. Após a obtenção do conjunto inicial de genes diferencialmente expressos é necessária a realização de diversos testes visando entender melhor a função geral dessas proteínas e sua importância para o processo de diferenciação. Isso é ainda mais relevante em *T. cruzi*, pois este é um organismo peculiar e relativamente distante dos principais organismos modelos eucarióticos, fatores que determinam um desconhecimento relativo de sua biologia molecular.

Nesse sentido, o presente trabalho representa a primeira abordagem comprobatória dentro dessa mudança de paradigma de avaliação dos processos biológicos, apresentando diversas características próprias que serão mais bem definidas nesse capítulo.

5.2 Genes selecionados a partir das análises de microarranjo de DNA

Conforme descrito no item 4.5 foram selecionados sete genes diferencialmente expressos, a partir dos dados de microarranjo de DNA, que foram organizados e selecionados de acordo com o perfil de hibridização e características funcionais dos genes em questão, conforme os critérios relatados anteriormente. Todos os genes principais selecionados apresentaram ortólogos e domínios definidos pelo banco de dados PFAM, pois estes foram critérios de seleção. A região gênica da qual a sonda é complementar encontra-se em sua grande maioria (seis dos sete genes selecionados), próxima à região 3' UTR do gene, o que indica uma boa confiabilidade com relação ao perfil de hibridização da sonda, devido à característica inerente ao processo de produção do material a ser hibridado, que tende a ser enriquecido em regiões mais próximas à cauda poli-A. Apenas dois dos genes relacionados foram identificados na análise do proteoma do *T. cruzi*, no trabalho publicado por ATWOOD *et al*, 2005. Embora a presença de tal dado represente uma maior confiabilidade na existência funcional dessa proteína, a ausência de identificação proteômica não constitui em evidência negativa forte, pois essa técnica apresenta, atualmente, uma sensibilidade baixa. A existência de ortólogos

em organismos relacionados, o que dificilmente ocorreria em regiões não-codificadoras, por outro lado, constitui uma evidência positiva da existência dessas proteínas.

Após esta primeira seleção, foram criados clusters para se selecionar novos genes, como citado no 4.5, utilizando-se critérios semelhantes. Na Tabela 5.1 estão listados todos os genes com suas nomenclaturas e principais características.

Tabela 5.1: Genes selecionados para caracterização.

ID IBMP	Classificação	Tamanho	Módulos (Domínios)	Proteômica	Ortólogos
TcExp02A01	Gene Principal	693 nt, 25 kDa	ALBA	Epi, Met, Ama, Tripo	Tb, Lm
TcExp02A08	Gene Acessório	1368 nt, 52 kDa			
TcExp02A09	Gene Secundário	1197 nt, 43 kDa	Saccharopine dehydrogenase	Meta	Lm
TcExp02A10	Gene Guia	1080 nt, 41 kDa	Sterol 24-c-methyltransferase, putative	Epi	Tb, Lm
TcExp02A02	Gene Principal	3126 nt, 117 kDa	Protein kinase		Tb, Lm
TcExp02A11	Gene Acessório	1644 nt, 61 kDa			
TcExp02A12	Gene Secundário	1320 nt, 49 kDa	Zinc finger C2H2-type)		Tb
TcExp02B01	Gene Guia	681 nt, 24 kDa	Ras-related protein rab-5, putative	Tripo	Tb, Lm
TcExp02A03	Gene Principal	3027 nt, 114 kDa	FKBP-type		Tb, Lm
TcExp02B02	Gene Acessório	1089 nt, 40 kDa			
TcExp02B03	Gene Secundário	954 nt, 35 kDa	Zinc finger, C2H2 type		Tb, Lm
TcExp02B04	Gene Guia	1095 nt, 41 kDa	DREV methyltransferase, putative		Tb, Lm
TcExp02A04	Gene Principal	1353 nt, 51 kDa	SWIM zinc finger		Tb
TcExp02B05	Gene Acessório	1485 nt, 53 kDa			
TcExp02B06	Gene Secundário	249 nt, 9 kDa	CSL zinc finger		Tb, Lm
TcExp02B07	Gene Guia	354 nt, 13 kDa	Dynein light chain 2B, putative		Tb, Lm
TcExp02A05	Gene Principal	606 nt, 22 kDa	ASF1-like		Tb, Lm
TcExp02B08	Gene Acessório	516 nt, 20 kDa			
TcExp02B09	Gene Secundário	1035 nt, 38 kDa	Programmed cell death protein 2		Tb, Lm
TcExp02B10	Gene Guia	531 nt, 20 kDa	Calcineurin B subunit, putative		Tb, Lm
TcExp02A06	Gene Principal	2247 nt, 86 kDa	Signal peptide/PPR Repeat		Tb, Lm
TcExp02B11	Gene Acessório	1905 nt, 72 kDa			
TcExp02B12	Gene Secundário	963 nt, 37 kDa	GTP-ase activating protein, Arf		Tb, Lm
TcExp02C01	Gene Guia	909 nt, 36 kDa	Syntaxin, putative		Tb, Lm
TcExp02A07	Gene Principal	1362 nt, 51 kDa	Kinase	Meta	Tb, Lm
TcExp02C02	Gene Acessório	1086 nt, 40 kDa	TLD		Tb, Lm
TcExp02C03	Gene Secundário	1209 nt, 48 kDa			Tb
TcExp02C04	Gene Guia	546 nt, 21 kDa	Centrin, putative		Tb, Lm

Pb (pares de bases); KDa (kilodalton); X (não encontrado); Lmj (Leishmania major); Tbr (Trypanosoma brucei).

5.3 Amplificação dos genes e purificação dos produtos

Dos 28 genes selecionados, todos foram amplificados com sucesso. Quanto à purificação dos produtos resultantes da PCR, houve perda de material durante o processo. Para resolver este problema, foi aumentado o número de ciclos na PCR de 30 para 35 visando o aumento da

quantidade dos produtos. Após essa modificação, obteve-se sucesso na purificação de todas as seqüências, conforme pode ser visto na Figura 5.1 para alguns genes selecionados. A facilidade com que essa etapa foi concluída reforça a qualidade do processo de predição dos iniciadores bem como a vantagem de se usar uma PCR com adaptadores externos, o que permite a utilização de uma temperatura de anelamento mais estridente nos ciclos mais tardios do processo de amplificação.

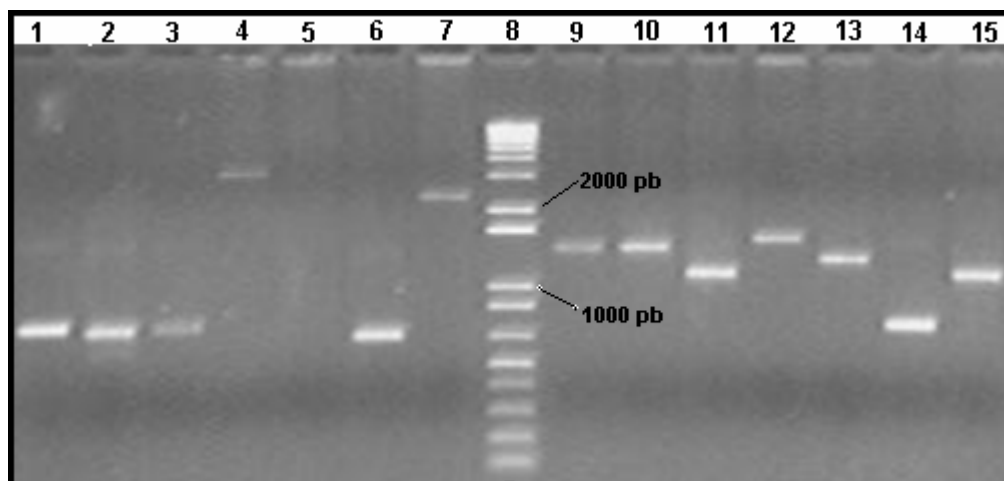


Figura 5.1 Padrão eletroforético obtido a partir das amostras de pcr purificadas.

As amostras numeradas de 1 a 7 e 9 a 15 representam o perfil eletroforético obtido dos produtos de pcr genes específicos contendo o sítio attB1 e 2. A amostra 8 é referente ao marcador de DNA(1 kb plus, Gibco), com dois indicativos de tamanho em pares de bases (pb).

5.4 Obtenção dos clones de entrada

A etapa de clonagem para obtenção dos clones de entrada foi realizada com sucesso, não necessitando qualquer alteração no protocolo descrito no capítulo 4. Na Figura 5.2 pode-se verificar o perfil eletroforético obtido na PCR de colônia para seleção dos clones positivos e na Figura 5.3 o perfil eletroforético dos plasmídeos contendo os insertos desejados já purificados conforme o protocolo descrito. Esse resultado demonstra a confiabilidade da plataforma escolhida para o desenvolvimento deste trabalho, bem como a facilidade na obtenção futura de um número muito maior de clones de entrada contendo seqüências do *T. cruzi*, com a intenção da construção do ORFeoma de *T. cruzi*.

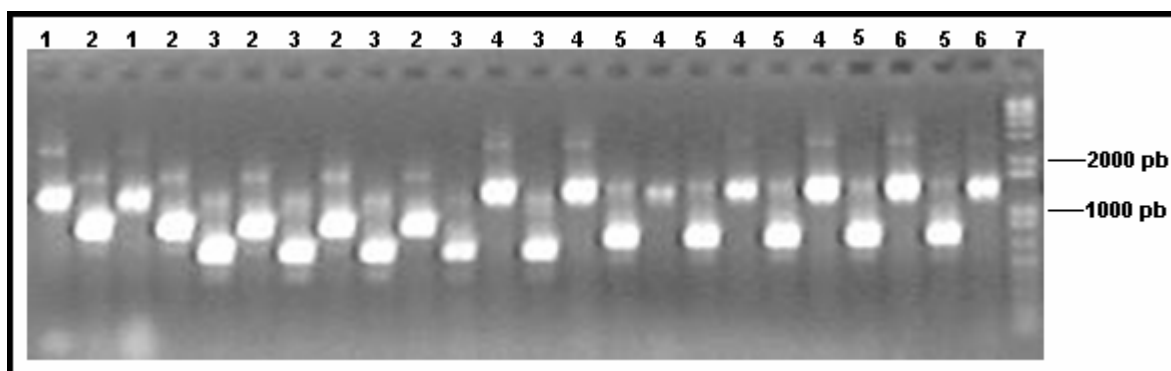


Figura 5.2 Perfil eletroforético dos produtos obtidos da PCR de colônia de clones de entrada.
A numeração das colunas de 1 a 6 exemplificam o perfil eletroforético da per de colônia para 6 distintos genes. A repetição das numerações é devida ao número de colônias testadas. A amostra 7 é referente ao marcador de DNA(1 kb plus, Gibco), com dois indicativos de tamanho em pares de bases (pb).

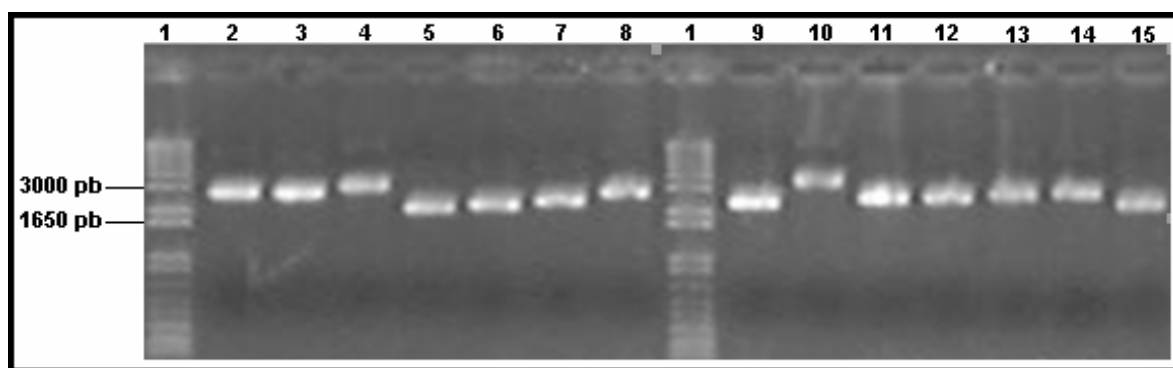


Figura 5.3 Perfil eletroforético dos clones de entrada purificados.
A numeração das colunas de 2 a 15 exemplificam o perfil eletroforético obtidos a partir dos clones de entrada contendo o gene de interesse para 13 distintos genes. A amostra 1 é referente ao marcador de DNA(1 kb plus, Gibco), com dois indicativos de tamanho em pares de bases (pb).

5.5 Obtenção dos clones de expressão

Da mesma forma que em relação aos clones de entrada, a etapa de clonagem para a obtenção dos clones de expressão foi realizada com sucesso, não necessitando qualquer alteração no protocolo descrito no capítulo 4. Na Figura 5.4 pode-se verificar o perfil eletroforético obtido na PCR de colônia para seleção dos clones positivos e na Figura 5.5 o perfil eletroforético dos plasmídeos de expressão (pDEST) contendo os insertos desejados já purificados conforme protocolo descrito. Estes resultados reforçam a eficácia do processo, essencial para a clonagem futura em outros vetores de destino, como por exemplo, os que estão sendo desenvolvidos no IBMP, com o intuito de obter informações em larga escala das proteínas de *T. cruzi*.

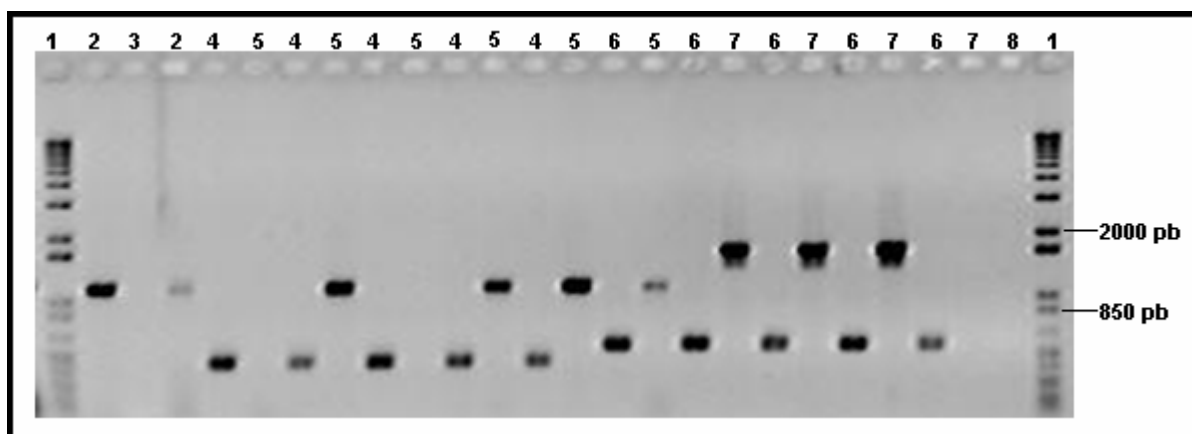


Figura 5.4 Perfil eletroforético dos produtos obtidos da PCR de colônia de clones de destinação.

A numeração das colunas de 1 a 8 exemplificam o perfil eletroforético da pcr de colônia de clones de destinação para 7 distintos genes. A repetição das numerações é devida ao número de colônias testadas. A amostra 1 é referente ao marcador de DNA(1 kb plus, Gibco), com dois indicativos de tamanho em pares de bases (pb).

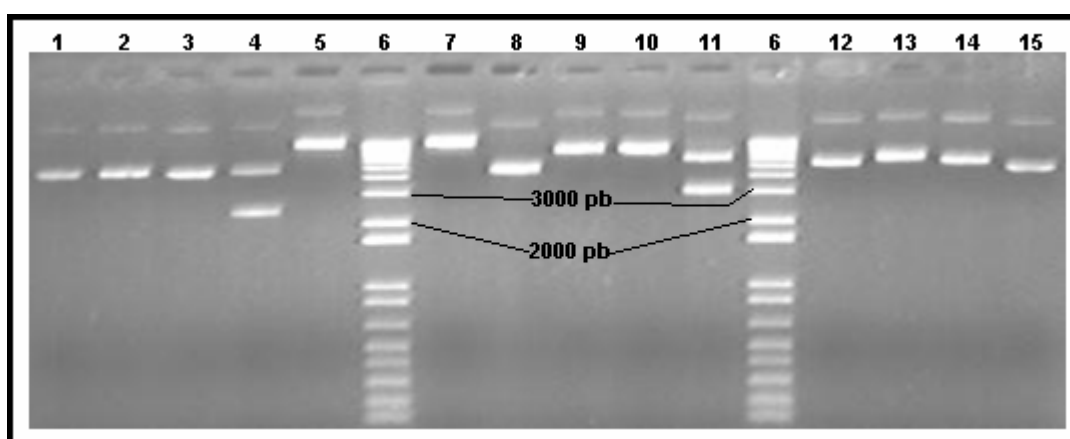


Figura 5.5 Perfil eletroforético dos clones de expressão purificados.

A numeração das colunas de 1 a 5 e 7 a 15 exemplificam o perfil eletroforético obtidos a partir dos clones de destinação contendo o gene de interesse para 14 distintos genes. A amostra 6 é referente ao marcador de DNA(1 kb plus, Gibco), com dois indicativos de tamanho em pares de bases (pb).

5.6 Seqüenciamento dos clones obtidos

As etapas de clonagem dos genes de interesse no vetor de entrada e de expressão obtiveram um grau de sucesso absoluto (100%). Porém, o gene 8728.t00021 (TcExp02B03) foi eliminado das etapas posteriores devido a um erro na construção do iniciador, pois foi identificada uma deleção de um nucleotídeo em sua seqüência, o que acarretou uma mudança de quadro de leitura na proteína recombinante expressa, sendo detectado pelos tamanhos dos fragmentos obtidos e por seqüenciamento, já que todos foram submetidos a este tipo de análise.

Através do seqüenciamento foi possível verificar e comprovar o grau de sucesso das clonagens sendo que todos os 27 clones apresentaram inserção correta e seqüência gênica esperada, nos clones de entrada e de expressão.

5.7 Expressão e purificação das proteínas

Na etapa de teste de expressão, a qual é realizada em um volume total de meio menor, foi obtida a expressão de 13 proteínas (48,1%), em uma quantidade minimamente aceitável, de acordo com uma avaliação subjetiva baseado na intensidade da coloração da banda da proteína expressa. Das 14 proteínas não selecionadas, quatro (14,8%) foram classificadas como de expressão duvidosa, nove (33,3%) como não expressas e uma (3,7%) como tendo tamanho diferente do esperado.

Dentre os nove clones nos quais a expressão protéica foi classificada como inexistente no teste, seis apresentavam um perfil minimamente sugestivo da existência de expressão em baixíssima quantidade. Na tentativa de obtenção da expressão desses seis genes, a temperatura durante a indução na fase de produção foi diminuída de 37°C para 22°C. Essa modificação do protocolo resultou na indução de expressão de duas proteínas (33,3% do total testado).

Na fase de produção, as 13 proteínas que haviam sido consideradas como expressas na fase de teste repetiram o mesmo resultado. Além disso, conforme mencionado acima, duas proteínas, classificadas como não expressas na fase de teste, tiveram seu protocolo de indução modificado na fase de produção, o que resultou na sua expressão. Finalmente, das quatro proteínas classificadas como tendo expressão duvidosa na fase de teste, duas foram expressas na fase de produção, após repetição do protocolo, o que sugere que razões técnicas pontuais causaram a falha na etapa de testes.

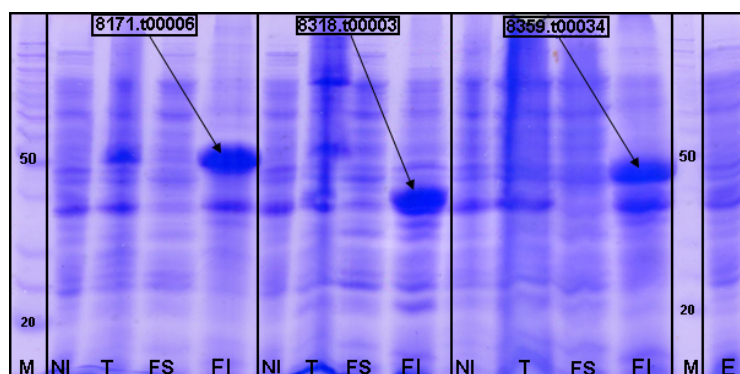
Portanto, foi obtida a expressão, na fase de produção, de 19 proteínas (70,4% do total inicial). Dos oito genes considerados como tendo uma expressão inexistente ou insuficiente para o prosseguimento do trabalho, quatro são codificadores de proteínas hipotéticas e quatro de proteínas hipotéticas conservadas. Dentre estas proteínas, somente uma (8694.t00004, TcExp02A08) apresenta peptídeos identificados pelo trabalho de proteômica (ATWOOD, 2005). Metade das proteínas não expressas apresentam alto peso molecular, sendo este um possível fator limitante na obtenção da expressão. Na Tabela 5.2 encontra-se o resultado das expressões das proteínas.

Tabela 5.2 Resultados da expressão de proteínas na fase de teste e produção.

Gene ID	Nomenclatura	Pb / KDa	Teste	Produção
8530.t00003	TcExp02A01	693 nt, 25 kDa	E	E
8694.t00004	TcExp02A08	1368 nt, 52 kDa	N	N
7626.t00009	TcExp02A09	1197 nt, 43 kDa	?	E
6680.t00001	TcExp02A10	1080 nt, 41 kDa	E	E
6996.t00063	TcExp02A02	3126 nt, 117 kDa	N	N
8773.t00014	TcExp02A11	1644 nt, 61 kDa	N	N
8364.t00002	TcExp02A12	1320 nt, 49 kDa	?	N
8208.t00006	TcExp02B01	681 nt, 24 kDa	N	E
7109.t00005	TcExp02A03	3027 nt, 114 kDa	N	N
8764.t00016	TcExp02B02	1089 nt, 40 kDa	N	E
8728.t00021	TcExp02B03	954 nt, 35 kDa	X	X
8257.t00019	TcExp02B04	1095 nt, 41 kDa	E	E
6896.t00027	TcExp02A04	1353 nt, 51 kDa	E	E
7224.t00006	TcExp02B05	1485 nt, 53 kDa	N	T
6958.t00014	TcExp02B06	249 nt, 9 kDa	?	E
8243.t00004	TcExp02B07	354 nt, 13 kDa	E	E
4715.t00001	TcExp02A05	606 nt, 22 kDa	T	E
5595.t00003	TcExp02B08	516 nt, 20 kDa	?	N
8171.t00006	TcExp02B09	1035 nt, 38 kDa	E	E
8416.t00006	TcExp02B10	531 nt, 20 kDa	E	E
8415.t00003	TcExp02A06	2247 nt, 86 kDa	N	N
7000.t00002	TcExp02B11	1905 nt, 72 kDa	N	N
8318.t00003	TcExp02B12	963 nt, 37 kDa	E	E
7168.t00007	TcExp02C01	909 nt, 36 kDa	E	E
5323.t00004	TcExp02A07	1362 nt, 51 kDa	E	E
8359.t00034	TcExp02C02	1086 nt, 40 kDa	E	E
7741.t00008	TcExp02C03	1209 nt, 48 kDa	E	E
6996.t00038	TcExp02C04	546 nt, 21 kDa	E	E

Pb (pares de base); KDa (kilodalton); E (proteína expressa); N (proteína sem expressão); T (tamanho não esperado); ? (não conclusivo), X (clone excluído) Em cinza, genes principais.

A avaliação da ocorrência de expressão protéica foi realizada através do uso de SDS-PAGE A Figura 5.6 mostra um padrão de expressão adequada bem como a qualidade obtida dos géis.

**Figura 5.6 Perfil eletroforético (SDS-PAGE) contendo o gene ID.**

M (marcador de peso molecular, kDa); NI (fração não induzida); T (fração total); FS (fração solúvel); FI (fração insolúvel); E (extrato controle pLysS).

Os 19 extratos contendo as proteínas recombinantes expressas foram então purificados conforme descrito no item 4.11, sendo que 18 proteínas foram obtidas, possibilitando a obtenção de anticorpos policlonais para cada uma delas. Uma das proteínas expressas foi

perdida durante o processo de purificação, devido a sua baixa expressão e necessidade de se realizar duas etapas de purificação devido à contaminação.

5.8 Obtenção dos anticorpos policlonais

Das 18 proteínas obtidas, todas foram inoculadas em camundongos em duplicata conforme descrito no item 4.12, sendo que ambos camundongos inoculados com a proteína referente ao gene 8257.t00019 (TcExp02B04) morreram, o mesmo tendo ocorrido com um camundongo inoculado com a proteína referente ao gene 8171.t00006 (TcExp02B09), obtendo-se, ao final, 33 soros (91,6%) dos 36 possíveis.

Utilizando a metodologia de *western blot*, os soros obtidos foram testados contra as proteínas recombinantes, sendo que todos os soros reconheceram suas respectivas proteínas. Contra os extratos de *T. cruzi*, do 17 anti-soros testados, 10 reconheceram suas proteínas (62,5% do total de anti-soros) com tamanho esperado, e um soro (6,25%), reconheceu especificamente a banda, porém com tamanho acima do esperado. As outras cinco proteínas (31,25%) não tiveram reação com o soro de modo satisfatório. Na Tabela 5.3 estão as proteínas que possuem anticorpos policlonais com qualidade adequada nos ensaios de *Western blot* utilizando extrato de *T. cruzi*.

Tabela 5.3 Tabela dos genes dos quais obteve-se anticorpo policlonal.

Nome interno	Qualidade	Função (Domínio)
TcExp02A01	Bom	Proteína Hipotética Conservada_ALBA
TcExp02A04	Bom	Proteína Hipotética Conservada_SWIM
TcExp02A09	Bom	Proteína Hipotética Conservada (Saccharopine dehydrogenase)
TcExp02A10	Bom	Sterol 24-c-methyltransferase, putative
TcExp02B01	Bom	Ras-related protein rab-5, putative
TcExp02B06	Químico	Proteína Hipotética Conservada (CSL zinc finger)
TcExp02B07	Bom	Dynein light chain 2B, cytoplasmic, putative
TcExp02B10	Bom	Calcineurin B subunit, putative
TcExp02B12	Bom	Proteína Hipotética Conservada (Putative GTP-ase activating)
TcExp02C01	Bom	Syntaxin, putative
TcExp02C04	Bom	Centrin, putative

Pb (pares de base); KDa (kilodalton); Químico (anticorpo policlonal com teste positivo apenas com quimioluminescência).

5.9 Imunolocalização por microscopia óptica

Conforme descrito no item 5.9 foram obtidos 11 soros policlonais com qualidade adequada para análise por *western blot* do padrão de expressão protéica durante a

metaciclógênese, sendo que esses mesmos soros foram utilizados para os ensaios de imunolocalização. Dos 11 soros testados, dois (TcExp02A04 e TcExp02B06) obtiveram um padrão inadequado para análise, pois houve reação cruzada com o soro pré-imune, o que foi identificado somente nessa etapa, sendo que ambos não apresentaram reação do soro pré-imune na fase anterior.

5.10 Características gerais dos genes selecionados e resultados de western blot e imunolocalização.

As características gerais dos genes como, por exemplo, aspectos gerais, homologias, funções dentre outros aspectos, serão abordadas nos tópicos a seguir bem como os resultados de *western blot* e imunolocalização por microscopia óptica.

5.10.1 TcExp02A01, 8530.t00003, Proteína hipotética conservada

Este gene é o principal do grupo 1 e possui em *T. cruzi* 616 pares de bases codificando uma proteína hipotética conservada de 25 kDa, apresentando um domínio PFAM denominado ALBA (*Acetylation Lowers Binding Affinity*). A proteína na qual esse domínio foi identificado inicialmente está presente em arqueobactérias possuindo aproximadamente 10 kDa e tem como função cobrir o DNA sem descompactá-lo (BELL *et al.*, 2002; SANDMAN e REEVE, 2005).

Este domínio foi encontrado primeiramente em estudos de proteínas de arqueobactérias sendo mais tarde denominada de ALBA, havendo diversos sinônimos: DBNP-B, Sac10b, Sso10b e Ssh10b. Esta proteína foi inicialmente caracterizada em algumas espécies de *Sulfolobus*, possuindo 10 kDa e tendo sua função regulada através da desacetilação por Sir2 e acetilação por PAT, regulando assim sua função mais conhecida, a de se ligar a fita dupla de DNA (MARSH *et al.*, 2005; VICTORIA *et al.*, 2005).

Além de se ligar à fita dupla de DNA, acredita-se também que a proteína em questão tenha capacidade de ligar-se ao RNA podendo estar associado com metabolismo e estabilidade do RNA (BOHRMANN *et al.*, 1994; ARAVIND *et al.*, 2003; GUO *et al.*, 2003; MARSH *et al.*, 2005).

O domínio ALBA apresenta uma conservação de vários aminoácidos como se pode verificar na Figura 5.7. O gráfico representa o grau de conservação dos aminoácidos nas posições formadoras do domínio em questão. No eixo X, a numeração é referente ao número

da posição aminoacídica e no eixo Y, ao grau de conservação do aminoácido em questão. Por exemplo, na posição 40 (eixo X) existe a possibilidade de ocorrência de glicina e aspartato, sendo que a glicina (letra G em maior tamanho) é muito mais conservada. Desta forma, pode-se verificar outras regiões conservadas.

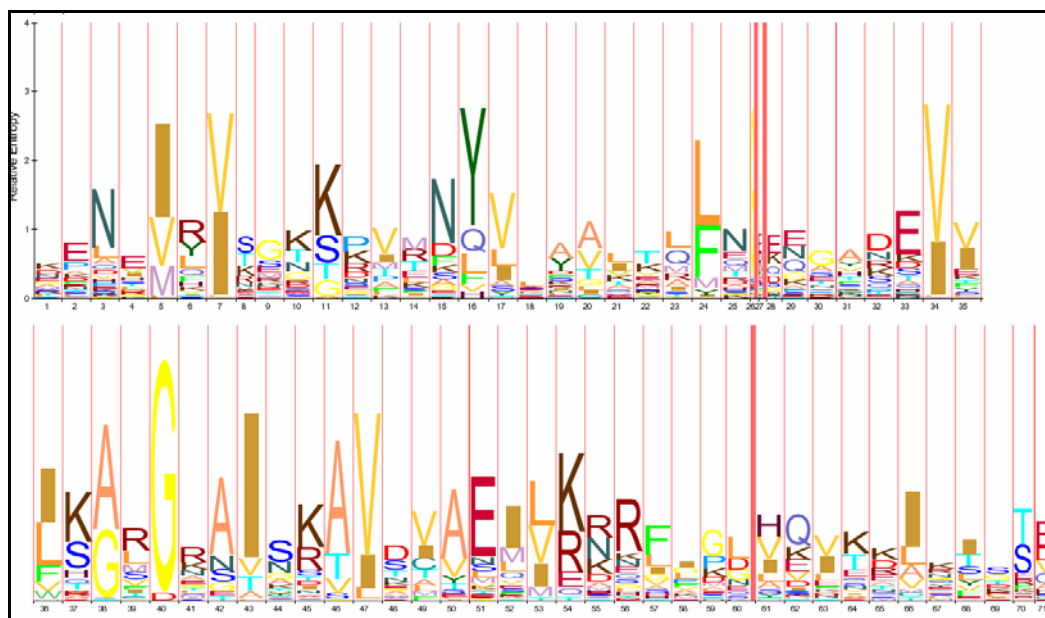


Figura 5.7 Gráfico LogoPlot representando o padrão de conservação do domínio quinase.

Representação do domínio ALBA no banco de dados PFAM, mostrando o grau de conservação em cada posição aminoacídica. No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa, onde a altura da pilha das letras em cada posição representa entropia relativa da distribuição das emissões dentro de algumas probabilidades em relação à distribuição de background dado por todo o perfil. A dimensão relativa de uma letra exprime a probabilidade de presença na posição a partir de um estado de distribuição. Em outras palavras: quanto maior a letra, maiores as informações sobre a respectiva posição nas famílias protéicas. As letras são ordenadas por ordem decrescente em função da sua probabilidade.

A confiabilidade da identificação do domínio PFAM ALBA na proteína de *T. cruzi* é alta ($1e^{-19}$) e o padrão de similaridade pode ser visto na Figura 5.8. No genoma do *T. cruzi* existem ainda mais duas proteínas que possuem o mesmo domínio, também classificadas como proteínas hipotéticas, mas apresentam um peso molecular menor, similar ao gene identificado em arqueobactérias: 4859.t00002/4903.t00007 (13 kDa) e 4859.t00001/4903.t00006 (13,5 kDa). Interessantemente, esses dois genes estão localizados justapostos no genoma de *T. cruzi*, mas são relativamente distintos entre si, pois mesmo sendo praticamente compostos somente pelo domínio ALBA, apresentam somente 51% de aminoácidos idênticos.

```

ALBA>keNeIrVsgkKpvmNYVlaaltlLnG.e.eegadeVvikArGrAIsk
eNeIrV+ k + Y++a+ lL+G+ ++ + d + i+A+G+AI
T. cruzi 26 PENEIRVTAKHGQRSYITYAIALLRGeDgKVQNDTIKISAMGAAIHN 72

ALBA> VdvAEilKrRfiglhqvkkikisteeitd<
AV++AEi+KrR++glhq +++ + i d
T. cruzi 73 AVNIAEIVKRRVVLHQTDDVSSE--IIHD 100

```

Figura 5.8. Comparação entre TcExp02A01 e o domínio ALBA.

Comparação utilizando o programa PFAM. Entre as duas seqüências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 26 da seqüência do *T. cruzi* e o término na posição 100.

A comparação da seqüência de *T. cruzi* com o proteoma de organismos modelos (Figura 5.9) identificou proteínas potencialmente ortólogas em eucariotos superiores, cuja similaridade maior se concentra na região do domínio ALBA, mas apresentando diversos aminoácidos conservados localizados em regiões externas ao domínio.

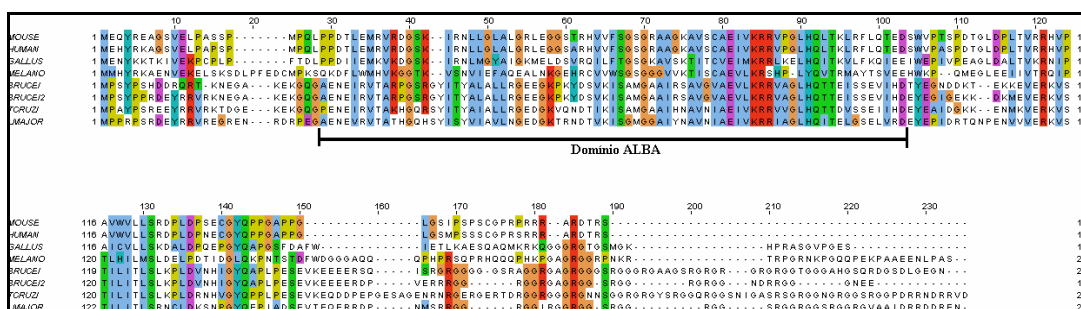


Figura 5.9 Análise de similaridade entre TcExp02A01 e ortólogos de outros eucariotos.

O Domínio ALBA está sendo demarcado abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

É possível verificar na Figura 5.9 a existência de uma região rica em repetições da seqüência RGG, somente em tripanossomatídeos, não sendo encontrada nos demais organismos representados. Esse motivo está associado à interação ao RNA (McBRIDE *et al*, 2005). reforçando o aspecto funcional de associação ao RNA de TcExp02A01. O alinhamento dos ortólogos de TcExp02A01 em tripanossomatídeos e outros protistas pode ser visto na Figura 5.10.

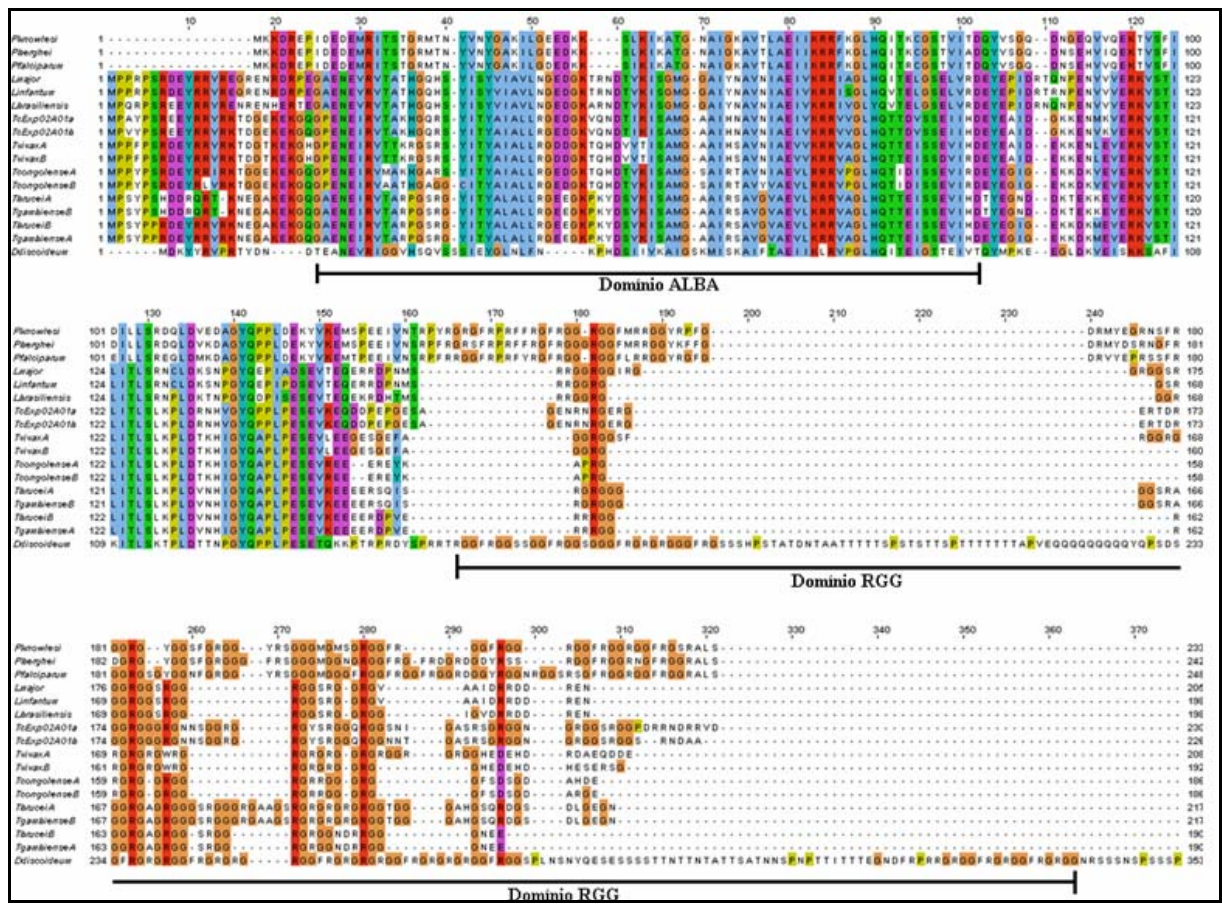


Figura 5.10 Análise de similaridade entre TcExp02A01 e ortólogos de protozoários.

Os Domínios ALBA e RGG estão sendo demarcados abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Em *T. cruzi*, *T. vivax*, *T. congolense*, *T. gambiense* e *T. brucei*, existem dois parálogos, contíguos no genoma, que diferem entre si na sua extremidade carbóxi-terminal. No gênero *Leishmania*, há somente um ortólogo. Também foi identificado um ortólogo no gênero *Plasmodium* e em *Dictyostelium discoideum*. A árvore filogenética dessas proteínas pode ser vista na Figura 5.11.

Ao se fazer uma busca pelo domínio ALBA e a existência de motivos RGG contíguos, foram identificados genes somente nas espécies supra-citadas, em *Arabidopsis thaliana* (dois genes) e mais um gene em *Plasmodium falciparum*, não tendo sido encontrado genes com esse padrão em eucariotos superiores. Interessantemente, fungos não apresentam proteínas com o domínio ALBA.

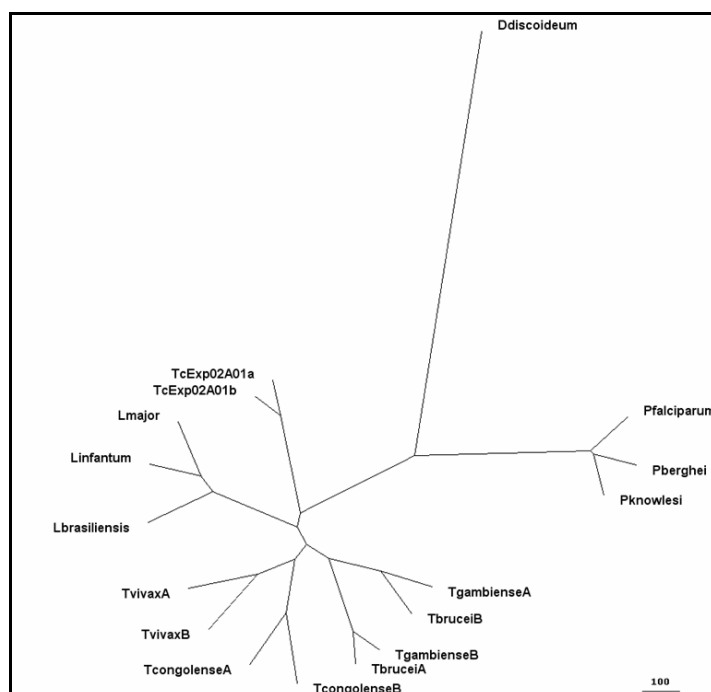


Figura 5.11 Árvore filogenética das proteínas contendo os domínios ALBA e RGG-box.
Quanto maior a distância entre as diferentes espécies menor a conservação das seqüências analisadas.

Dentro do fluxograma do presente trabalho, obteve-se sucesso para todas as etapas. O resultado da análise por *Western blot* da metaciclogênese pode ser visto na Figura 5.12, na qual observa-se claramente um padrão de diminuição durante a metaciclogênese, em relação a epimastigotas, com 1,4 vez em parasitas em diferenciação por 24 horas e 1,9 vez em tripomastigotas metacíclicas.

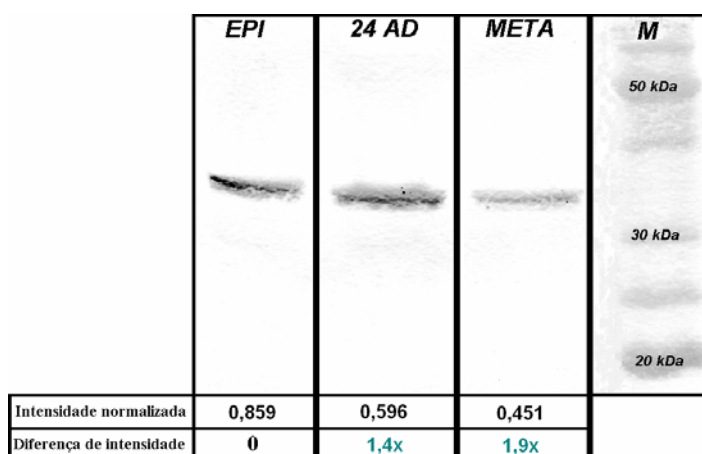


Figura 5.12 Análise de expressão da proteína TcExp02A01 por *western blot* em extratos de *T. cruzi*.
EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em verde representam diminuição da expressão em relação a epimastigotas.

A imunolocalização de TcExp02A01 apresentou um padrão granular por todo citoplasma celular como pode ser observado na Figura 5.13.

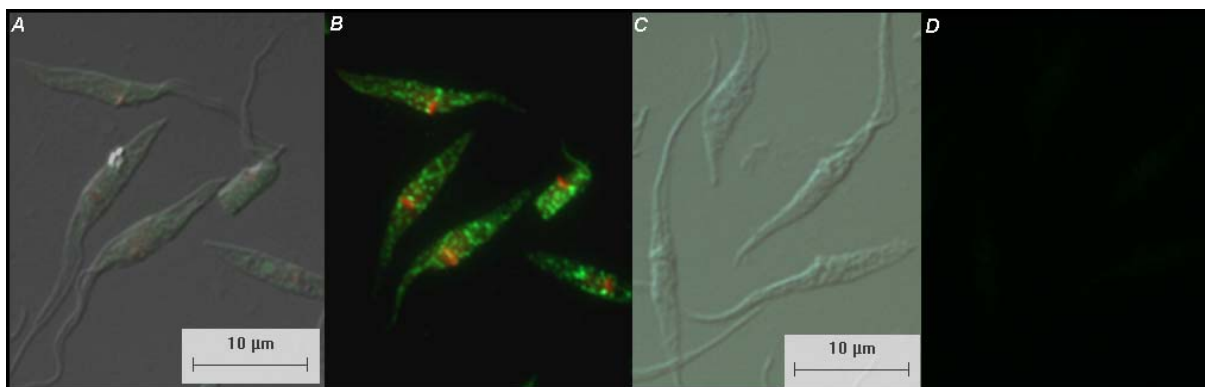


Figura 5.13 Imunolocalização de TcExp02A01.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e marcação do DNA (DAPI, vermelho); C) contraste diferencial de fase do soro pré-imune. D) Imunofluorescência do soro pré-imune.

5.10.2 TcExp02A08, 8694.t00004, Proteína Hipotética

Este gene é o acessório do grupo 1 e possui em *T. cruzi* 1368 pares de bases codificando uma proteína hipotética de 52,1 kDa, não apresentando domínios nos bancos de dados PFAM e Interpro. Essa proteína foi anotada inicialmente como hipotética, isto é, além de não ter sua função atribuível, também não havia sido encontrado ortólogos em outros organismos. No entanto, com o avanço do sequenciamento de outros tripanossomatídeos, foi possível identificar um possível ortólogo em *L. infantum*, denominado LinJ31_V3.2550, com aproximadamente 47% de similaridade. Portanto, sua anotação deve ser modificada para proteína hipotética conservada.

Interessantemente, há uma proteína de *T. cruzi*, 7963.t00006, com um grau de similaridade suficiente (37%) com TcExp02A08 para correlacionar essas duas proteínas entre si. Esse outro gene, anotado como proteína hipotética conservada, possui um peso molecular estimado de 65,5 kDa, não apresentando domínios protéicos identificados, e estando presente em um número maior de tripanossomatídeos, incluindo *L. major* mas excluindo *T. brucei*. A comparação entre TcExp02A08 e 7963.t00006 pode ser vista na Figura 5.14, no qual evidencia-se que, embora relacionadas, apresentam divergência suficiente para serem consideradas distintas evolutiva e funcionalmente.

É importante ressaltar que para ambos genes de *T. cruzi* só foi identificado ortologia com o gênero *Leishmania*, uma possível indicação de adaptação funcional dessa proteína.

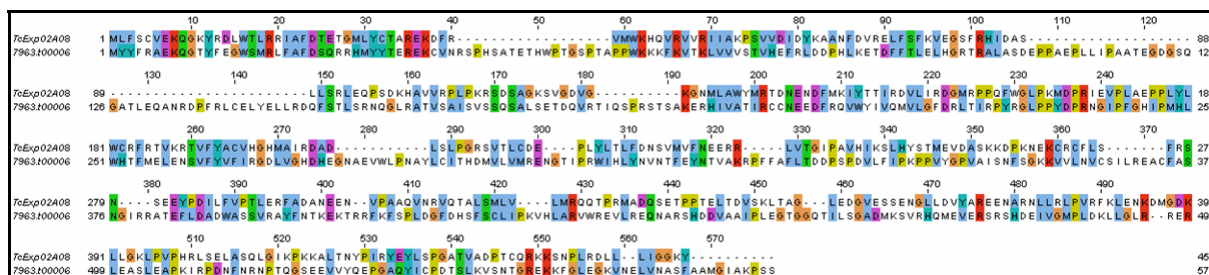


Figura 5.14 Comparação entre TcExp02A08 e 7963.t00006.

Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado, obtendo-se o vetor de entrada e de expressão. Porém não foi obtido sucesso na etapa de expressão em *E. coli*.

5.10.3 TcExp02A09, 7626.t00009, Proteína hipotética conservada

Este gene está classificado como gene acessório do grupo 1 e possui em *T. cruzi* 1197 pares de bases codificando uma proteína hipotética conservada de 42,9 kDa, apresentando um domínio denominado sacaropina desidrogenase ([N6-(glutaryl-2)-l-lysine:NAD oxidoreductase (l-lysine forming)]). Na anotação original do genoma, esse domínio não foi identificado. Em nossa análise apresentou um valor de confiabilidade de $1,4 \times 10^{-17}$, e seu ortólogo em *L. major* tem a anotação desse domínio, representando, portanto, uma falha na anotação desse gene pelo consórcio de seqüenciamento. O domínio se estende por praticamente toda a proteína, da posição 9 à 384, de um total de 399 aminoácidos, conforme pode ser verificado na Figura 5.15. Não foram identificadas outras proteínas de *T. cruzi* com esse domínio.

Esta enzima tem como função catalizar a etapa final da biossíntese de lisina (XU *et al*, 2007). Ao se comparar a seqüência protéica de TcExp02A09 com o proteoma de tripanossomatídeos, foi possível evidenciar ortólogos em *L. major* (LmjF28.0380), *L. infantum* (LinJ28_V3.0520) e *T. vivax* (tviv1669g12.p1k_6), o qual, atualmente, apresenta um genoma parcialmente seqüenciado, não sendo identificado ortólogo em *T. brucei* e outras Leishmanias. O alinhamento dos quatro ortólogos está representado na Figura 5.16 sendo possível observar o grande grau de conservação de aminoácidos similares entre essas quatro proteínas.



Figura 5.15 Comparação entre TcExp02A09 e o domínio sacaropina desidrogenase.

Entre as duas seqüências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 9 da seqüência do *T. cruzi* e o término na posição 384.

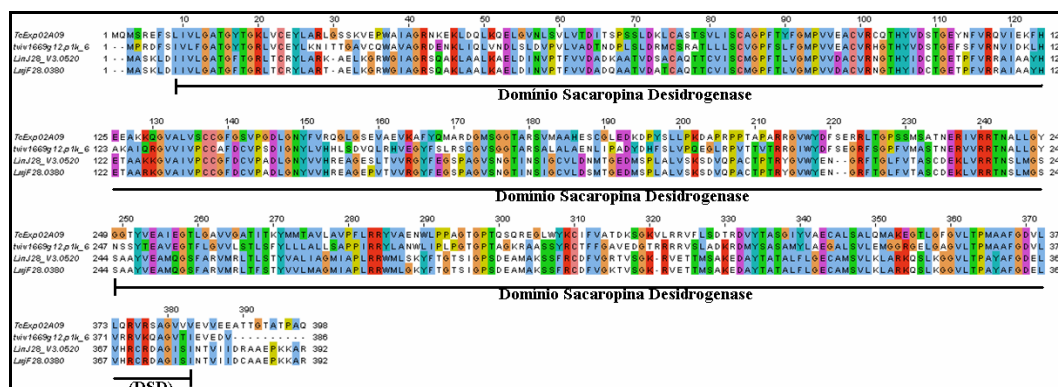


Figura 5.16 Comparação entre a proteína codificada por TcExp02A09 e os ortólogos encontrados.

O domínio sacaropina desidrogenase está delimitado abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Ao se ampliar a comparação a outros organismos, observa-se um grau de similaridade relativamente alto com um grande número de organismos das mais diferentes divisões filogenéticas, sendo que os mais próximos são proteobactérias (50% de similaridade em uma região de 413 nucleotídeos, abrangendo toda a proteína), mas tendo um grau semelhante de similaridade, em uma região menor compreendendo aproximadamente 150 aminoácidos iniciais, com outros organismos das mais diferentes divisões filogenéticas.

O alinhamento de TcExp02A09 com uma proteína mais similar em bactéria e em mamíferos pode ser vista na Figura 5.17, sendo respectivamente uma cianobactéria, *Gloeobacter*

violaceus (e-value de $6e^{-52}$, 50% de similaridade), e chimpanzé, *Pan troglodytes* (e-value de $3e^{-32}$, 46% de similaridade).

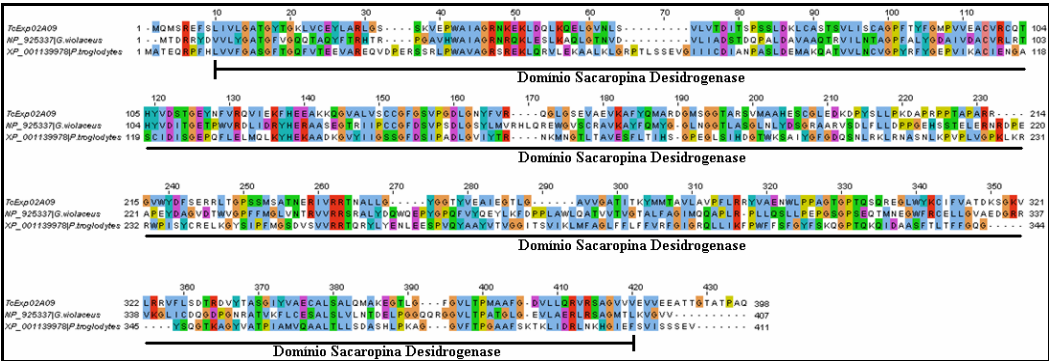


Figura 5.17 Análise de similaridade entre TcExp02A09 e os ortólogos em *G. violaceus* e *P. troglodytes*. O domínio sacaropina desidrogenase está delimitado. Na parte superior encontra-se a numeração relativa à posição aminoacídica.

As etapas do fluxograma de análise para este gene foram finalizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.18, na qual se observa claramente uma forte diminuição da expressão, em relação a epimastigotas, de 3,1 vezes em aderido 24 horas e de 1,5 vez em metacíclico. O anti-soro apresentou sinal em uma faixa de peso molecular não esperado, próxima a 25 kDa, nas forma epimastigota e aderido 24 horas.

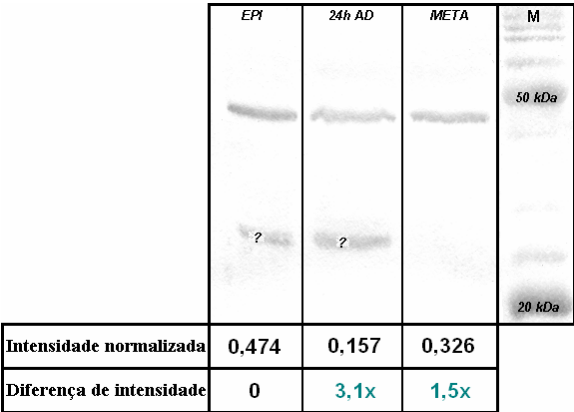


Figura 5.18. Análise de expressão da proteína TcExp02A09 por *western blot* em extratos de *T. cruzi*. EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular; ?: suposta degradação. Os valores em verde representam diminuição da expressão em a epimastigotas.

É importante ressaltar, conforme mencionado acima, que TcExp02A09 é a única proteína em *T. cruzi* que contém esse domínio e, portanto, a possibilidade de reação cruzada é remota,

não tendo sido evidenciado, também, indícios de degradação protéica. Uma hipótese a ser testada é se essa proteína de ~25 kDa representa uma forma truncada de TcExp02A09.

A imunolocalização da proteína TcExp02A09 apresenta um padrão de imunofluorescência difuso por todo citoplasma e também uma região de maior intensidade no cinetoplasto, melhor evidenciada com um menor tempo de exposição (C) como pode ser verificado na Figura 5.19.

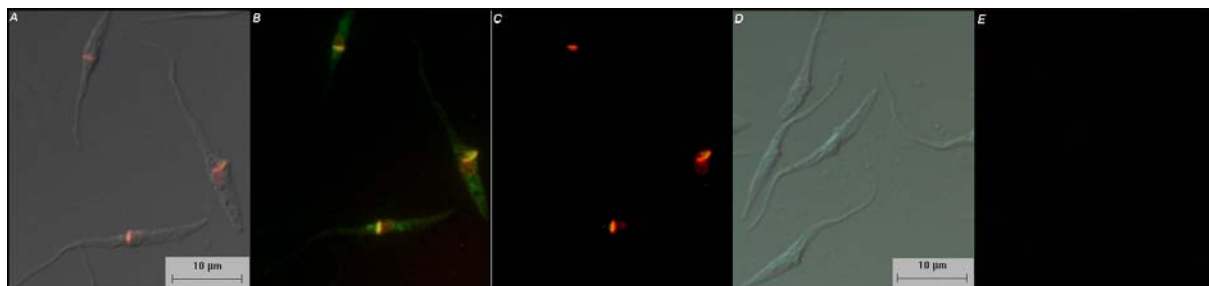


Figura 5.19 Imunolocalização da proteína TcExp02A09.

A) contraste diferencial de fase com sobreposição de C. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e marcação do DNA (DAPI, vermelho) C) imunolocalização com menos exposição; D) contraste diferencial de fase do soro pré-imune. E) Imunofluorescência do soro pré-imune.

5.10.4 TcExp02A10, 6680.t00001, Sterol 24-c-methyltransferase

Este gene é o gene guia do grupo 1 e possui em *T. cruzi* 1080 pares de bases codificando uma proteína hipotética conservada de 40,5 kDa, a qual apresenta um domínio denominado Sterol_MT_C (esterol metiltransferase C-terminal), sendo esta proteína a única em *T. cruzi* com esse domínio. Este domínio é encontrado principalmente na região carboxi-terminal de metiltransferases de fungos e plantas envolvidas com a metilação de esteróis (BOUVIER-NAVE *et al*, 1998).

A comparação da região da proteína de *T. cruzi* com o domínio pode ser vista na Figura 5.20. Além de apresentar similaridade com esse domínio, entre as posições 216 e 351 (valor de confiabilidade $2e^{-55}$), também apresenta similaridade com o domínio Methyltransf11 (metiltransferase, entre as posições 109 e 207, valor de confiabilidade $5e^{-21}$), o qual é encontrado em proteínas que transferem grupamentos metil a partir de S-adenosil-L-metionina (SAM) para átomos de nitrogênio, oxigênio ou carbono de DNA, RNA, proteínas ou moléculas pequenas. Essas modificações apresentam funções regulatórias, estando envolvida em processos de regulação da expressão gênica e diferenciação (PAWLAK *et al*, 2000).

Esses dois domínios, se estendendo por praticamente toda a proteína de 359 aminoácidos, são compatíveis entre si e reforçam a confiabilidade da predição. No entanto, a anotação do genoma pelo consórcio de sequenciamento não identificou tais domínios, evidenciando mais uma falha no processo, a qual, nesse caso, é menos problemático, pois a anotação funcional da proteína a identifica corretamente como uma metiltransferase.

```

Alignment of Sterol_MT_C vs T. cruzi /216-351

Sterol_MT_C->YDpnNpeHreIaygIElGDGIPkmtvkealeAlKkVGF.....Ev
YDpnN+eHr+I++ IElGD++P+++t ++++eAlK GF +++ +v
- T. cruzi 216 YDPNNEEHRRIRHKIElGDSLPLETKGQVIEALKASGFivedsfDV 262

Sterol_MT_C lheedLAddDplPWYyPLdGewkyvqslwDsltifRmtriGRwithnmV
+ ++ ++ lPWY+ L G+ + l +++ ++GRw+t+ m+
- T. cruzi 263 AERFESSPIH-NLPWYLTlQGN-----YTTLAGLKCSPLGRWFTNKMC 304

Sterol_MT_C gvLEkiGLAPKsKKVtdaLekAAegLVaGGkkklFTPMfLfvaRKP<-
+vLE+ GLAP G+ k +++Le+AA +LV GG+ ++FTP f+ aRKP
- T. cruzi 305 RVLEFAGLAPSGTHKGAEILEEAARSLVSGGESGIFTSPFFAKARKP 351

Alignment of Methyltransf_11 vs - T. cruzi

Methyltransf_11->LDvGcGtGllaraLarrvgpgarvtGvDlspemlalArerapragle
LD+GcG G ar + r+ g++v Gv +++++ +Ar++ + g +
- T. cruzi 109 LDLGCGVGGPARNIVRL--AGCNVMGVNNNEYQISRARRHDTKYGMN 153

Methyltransf_11 n...fvvgDaedLPfpDesFDlVsslvllhlaedperalrEiaRVLKPG
+ ++ + D++++ f D+ +D +++++ +h +d ++++E++RV KPG
- T. cruzi 154 SkinYTKTDFCNMCFGDNEYDGAYAIEATCHA-TDKVKCFSEVERVIKPG 202

Methyltransf_11 Gklvi<-
++v+
- T. cruzi 203 SYFVL 207

```

Figura 5.20 Comparação entre TcExp02A10 e o domínio Sterol_MT_C e Methyltransf_11.

Entre as duas seqüências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 216 e o término na posição 351 com relação ao domínio Sterol_MT_C, e início na posição 109 e término em 207 para o domínio Methyltransf_11.

A análise de TcExp02A10 com outros eucariotos organizados no banco de dados Homologene pode ser vista na Figura 5.21. As regiões mais conservadas estão localizadas dentro dos domínios, sendo que o domínio Methyltransf11 é mais conservado do que o Sterol_MT_C. Os aminoácidos entre as posições 77 e 127 também apresentam um alto grau de conservação, embora essa região não apresenta um domínio identificado. Esse último dado sugere que as proteínas analisadas são ortólogas entre si e que tal região apresenta importância funcional ou estrutural.

Baseado nestas análises pode-se verificar um padrão adequado de conservação do domínio, existindo uma região inicial com menor similaridade entre as seqüências até a posição aminoacídica 70, uma região mais conservada (entre 70 até 126) seguida do domínio Methyltransf11 (132 até 232) e do domínio Sterol_MT_C (240 até 286) respectivamente.

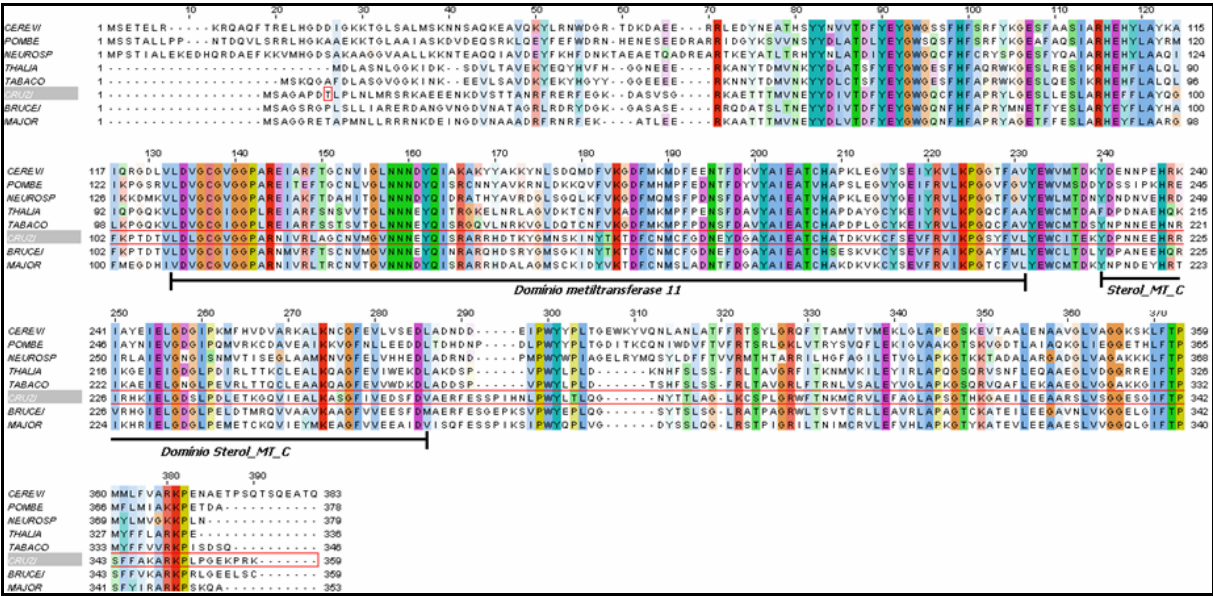


Figura 5.21 Análise de similaridade entre seqüências de *T. cruzi* e outros eucariotos. Os Domínios Metiltransferase 11 e Sterol_MT_C estão demarcados abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

As etapas do fluxograma de análise para este gene foram finalizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.22, na qual se observa diminuição da expressão em relação a epimastigota de 2,1 vezes em aderido 24 horas e de 31 vezes em metacíclicos.

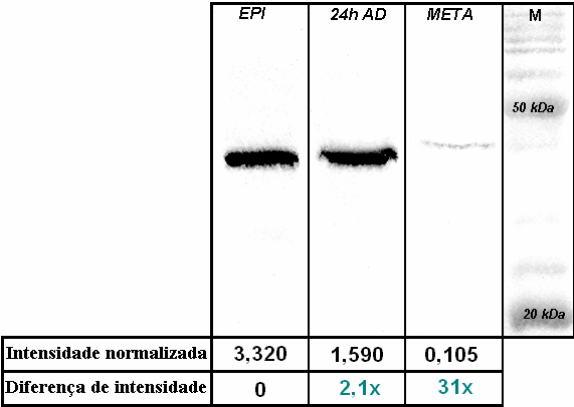


Figura 5.22. Análise de expressão da proteína TcExp02A10 por *western blot* em extratos de *T. cruzi*. EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em verde representam diminuição da expressão em relação a epimastigotas

A imunolocalização da proteína TcExp02A10 apresentou um padrão difuso pelo citoplasma, sem a existência de agrupamentos distinguíveis, como pode ser verificado na Figura 5.23.

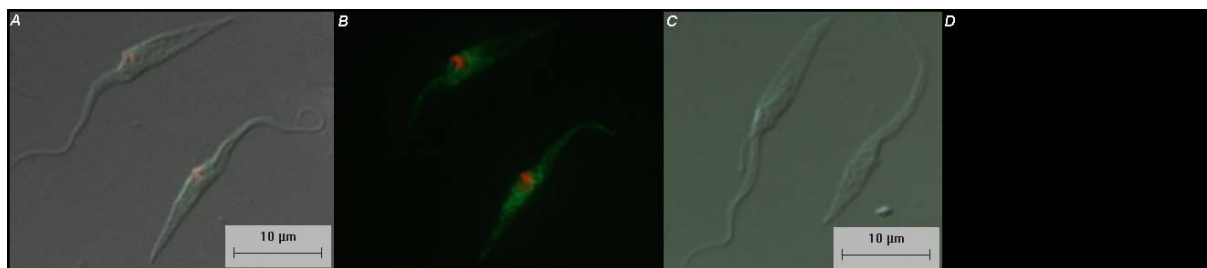


Figura 5.23 Imunolocalização da proteína TcExp02A10.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal anti-6680.t00001. Em verde a marcação referente ao anticorpo específico e em vermelho a marcação de DNA utilizando DAP. C) contraste diferencial de fase relativo ao soro pré-imune. D) Imunofluorescência obtida com a utilização do soro pré-imune.

5.10.5 TcExp02A02, 6996.t00063, Proteína Hipotética Conservada

Este gene está classificado como o gene principal do grupo 2 e possui 3196 pares de bases em *T. cruzi*, codificando uma proteína hipotética conservada de 117 kDa, a qual apresenta um domínio de proteína-quinase. Esse domínio faz parte de um extenso grupo de proteínas que compartilham um domínio catalítico conservado em serina/treonina e tirosinas quinases, apresentando diversas glicinas conservadas em sua região aminoterminal, próximo a um resíduo de lisina também conservado, o qual parece estar associado à ligação de ATP. Além destes resíduos, existem resíduos de ácido aspártico conservados na região central do domínio, os quais são importantes para a atividade catalítica da enzima. Estes resíduos estão evidenciados na Figura 5.24. Essa família de proteínas está presente em eucariotos, vírus e algumas bactérias.

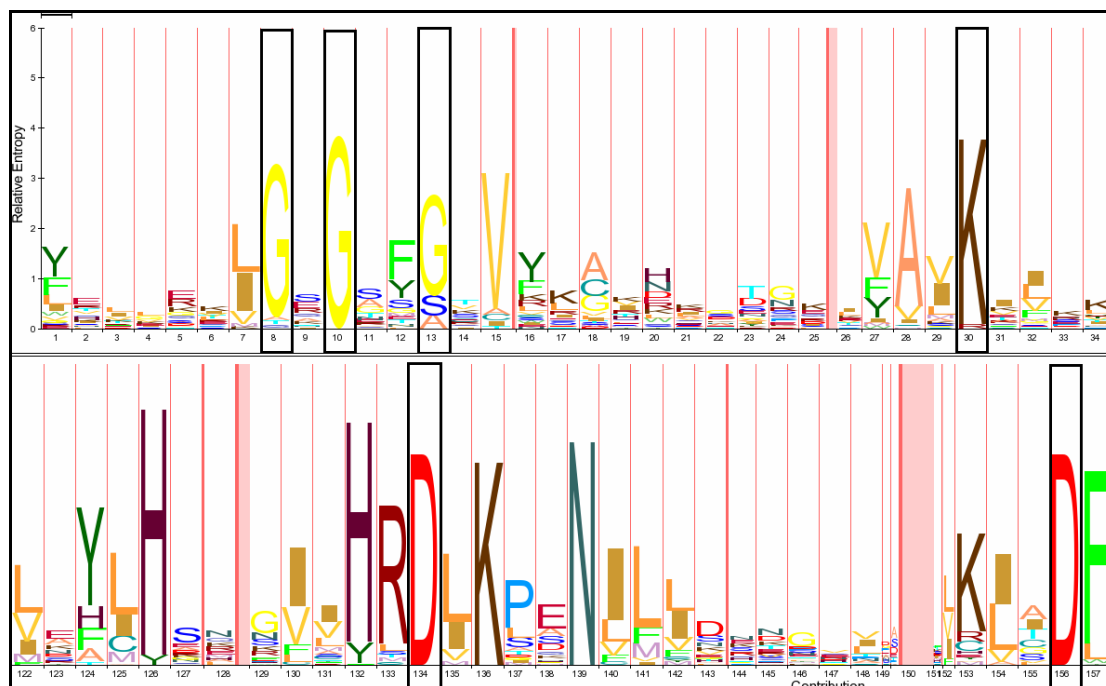


Figura 5.24 Gráfico LogoPlot representando o padrão de conservação do domínio quinase.

Representação do domínio quinase no banco de dados PFAM, mostrando o grau de conservação em cada posição aminoacídica. No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa. Os aminoácidos G (glicina), K (lisina) e D (ácido aspártico) característicos do domínio estão demarcados em retângulos pretos.

Quando comparada a similaridade entre o domínio PKinase e a sequência do *T. cruzi*, verificou-se um valor de confiabilidade de $7,5e^{-35}$. O início do domínio na sequência protéica de TcExp02A02 ocorre na posição 666 e o seu término ocorre na posição 972 como mostra a Figura 5.25. É interessante notar que embora essa proteína apresente 1041 aminoácidos, somente 307 (29,5%) abrangem esse domínio. Como este domínio está localizado na extremidade carboxi-terminal, temos uma região relativamente ampla sem a evidência de qualquer outro domínio em sua porção amino-terminal. Geralmente, proteína-quinases apresentam domínios acessórios, que auxiliam na determinação de sua especificidade. A ausência desses domínios acessórios em TcExp02A02, associado à existência de uma região protéica grande sem domínios identificados, pode ser indício de que haja um domínio acessório novo, uma situação relativamente comum em proteína-quinases de *T. cruzi* (EL-SAYED et al, 2005a).



Figura 5.25. Comparação entre a sequência de *T. cruzi* e o domínio PKinase.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 666 da sequência do *T. cruzi* e o término na posição 972.

A identificação dos ortólogos em tripanossomatídeos foi realizada, tendo-se encontrado em todas as espécies sequenciadas até o momento. O alinhamento das sequências protéicas obtidas pode ser visto na Figura 5.26 e, de forma geral, podemos verificar que são bastante conservada, mesmo na região não abrangendo o domínio quinase. As proteínas de *T. vivax* e *T. congolense* são menores provavelmente porque estão incompletas, pois esses dois genomas ainda estão em fase de montagem. Quando comparamos essa sequência com outros organismos, é obtido um número muito grande de sequências cuja similaridade é ditada pela região contendo o domínio, o que impossibilita a clara definição de homologia.

Um resultado interessante que resulta da comparação de TcExp02A02 com as proteínas de outros organismos é a grande semelhança da região do domínio com proteínas anotadas como quinase 2 do fator de iniciação da tradução 2 (EIF2AK2). Devido a isto, a anotação dessa proteína em tripanossomatídeos é, em geral, derivada dessa similaridade. No entanto, conforme mencionado acima, a região de similaridade está contida no domínio, o que enfraquece a associação evolutiva. Interessantemente, as proteínas EIF2AK2 têm um tamanho muito próximo ao do domínio quinase e, portanto, a região anterior ao domínio da proteína de *T. cruzi* não poderia realmente apresentar similaridade com EIF2AK2. Uma possibilidade que pode se aventar é a de que TcExp02A02 é realmente o ortólogo de EIF2AK2 e sofreu fusão

Este gene foi clonado, obtendo-se o vetor de entrada e de expressão, porém não foi possível obter uma proteína expressa.

5.10.7 TcExp02A12, 8364.t00002, Proteína Hipotética conservada

Este gene é o secundário do grupo 2 e possui em *T. cruzi* 1320 pares de bases codificando uma proteína hipotética conservada de 48,9 kDa, a qual não apresenta domínio predito pelo programa PFAM, mas com a predição de um peptídeo sinal. Ao tentar se identificar ortólogos por similaridade, foram identificadas somente seqüências em *T. vivax*, *T. brucei*, *T. congolense* e *T. b. gambiense*, com e-value próximo a $1e10^{-15}$, um valor relativamente baixo. No entanto, há indicação de ortologia entre os genes de *T. cruzi* e *T. brucei* na análise disponível em <http://www.genedb.org>, a qual leva em consideração a conservação de sentença, além da similaridade de seqüência. Não foram identificados ortólogos em *Leishmania*, um dado importante. De maneira geral, a conservação entre a seqüência de *T. cruzi* e os outros tripanossomatídeos é baixa, como pode ser visto na Figura 5.27. Ela apresenta uma inserção, entre os aminoácidos 96 e 132. Interessantemente, a região hidrofóbica presente na seqüência de *T. cruzi* que caracterizaria o peptídeo sinal não é visto em nenhuma das outras seqüências de tripanossomatídeos e a seqüência de *T. cruzi* apresenta uma metionina interna, logo a jusante do peptídeo sinal, na posição 27. Esse padrão é indicativo de um erro de predição de início da região codificadora, pois é um padrão recorrente na montagem do genoma de *T. cruzi*.

Essa informação foi levada em conta na predição do iniciador *forward* utilizado no presente trabalho e, portanto, a proteína recombinante produzida no presente trabalho teria um peso molecular estimado de 45,7 kDa.

Em geral, cada gene apresenta mais de uma região codificadora identificada pelo projeto genoma de *T. cruzi*, causada principalmente pela natureza híbrida da cepa selecionada para o seqüenciamento. Muitas vezes, a predição do códon de início da tradução é diferente entre essas duas seqüências, sendo que a seqüência com um códon de iniciação mais a montante apresenta, muito freqüentemente, uma grande quantidade de timinas, a qual é traduzida para fenilalanina, o que dá o caráter hidrofóbico a essa região.

No entanto, ao se comparar a seqüência cujo códon de iniciação está localizado mais a jusante, ela apresenta uma região pré-ATG idêntica ao gene com a CDS maior, o que reforçará o erro de predição. Além disso, os ortólogos em *T. brucei* e *L. major*, genomas cuja montagem é melhor, não apresentam essa região hidrofóbica. É importante salientar que o sinal de trans-splicing, o qual está presente geralmente em uma região muito próxima ao ATG inicial, é rico em poli-pirimidinas.

Embora não tenham sido identificados domínios protéicos com um grau de confiabilidade aceitável, há a predição de um domínio com valor de confiabilidade acima do limiar, denominado zf-TRAF (PF02176), o qual é encontrado em algumas proteínas associadas com o receptor do fator de necrose tumoral, o qual não ocorre em *T. cruzi*, indicando haver um domínio *zinc-finger* degenerado em TcExp02A12.

Interessantemente, há diversos resíduos de histidina e cisteína conservados nos quatro tripanossomatídeos selecionados: C227, C230, C247, C254, H243, H270 e H278. Esses resultados reforçam a existência de um domínio relacionado aos de *zinc finger* ainda não bem caracterizado.

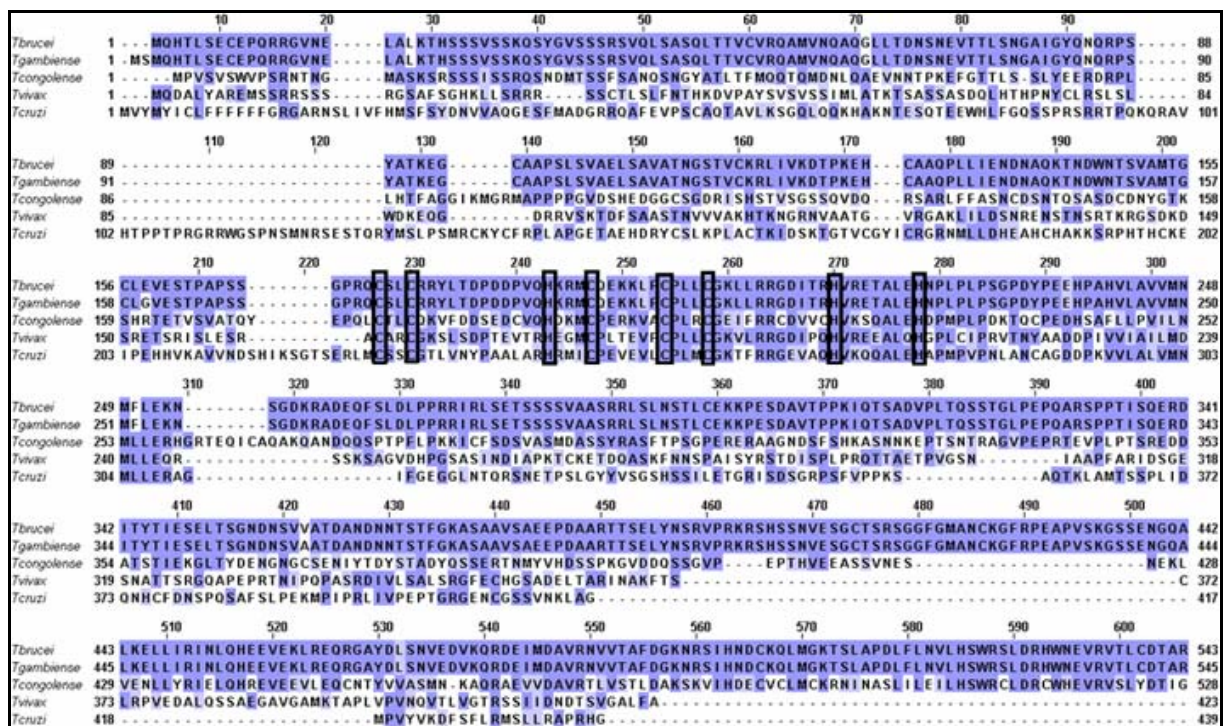


Figura 5.27 Análise de similaridade entre seqüências de *T. cruzi* e outros tripanossomatídeos.

Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoácida. As posições destacadas com retângulos pretos são relativos às histidinas e cisteínas conservadas descritas neste trabalho.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém não foi possível obter uma proteína expressa.

5.10.8 TcExp02B01, 8208.t00006, Ras-related protein rab-5

Este gene está classificado como guia do grupo 2 e possui em *T. cruzi* 681 pares de bases codificando uma proteína hipotética conservada de 24 kDa, a qual apresenta um domínio denominado RAS da superfamília de pequenas GTPases. Essa família de proteínas é bastante

numerosa contendo algumas sub-famílias principais, como RAB, RAS, RAC, RAL, dentre outras.

Ao analisar a sequência em *T. cruzi* no programa PFAM pode-se encontrar uma alta similaridade entre o domínio RAS e a sequência de *T. cruzi*, com valor de confiabilidade de $1e^{-74}$. O início do domínio na sequência protéica de *T. cruzi* ocorre na posição 17 e seu término na posição 190, como mostra a Figura 5.28, ocupando assim quase a totalidade da sequência aminoacídica.

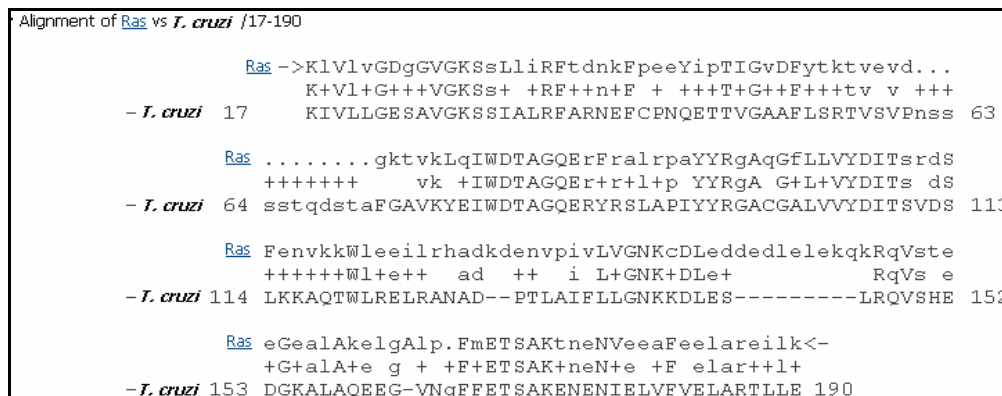


Figura 5.28. Comparação entre TcExp02B01 e o domínio RAS

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 17 da sequência do *T. cruzi* e o término na posição 190.

A análise da sequência de *T. cruzi* com o uso do programa BLASTp identificou diversos genes similares, anotados como Rab5, reforçando a anotação correta desse gene e a conservação grande desse domínio, de leveduras a humanos. A análise das sequências homólogas pode ser vista na Figura 5.29, verificando-se uma grande conservação do domínio, com a existência de uma região inicial com menor similaridade entre as sequências, seguida de regiões com maior similaridade, separadas por sequências específicas de tripanossomatídeos e de *A. thaliana* (posição 69 a 89).

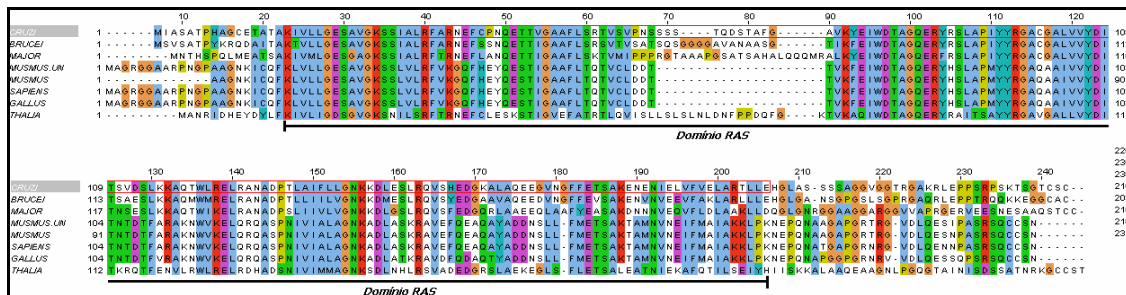


Figura 5.29 Análise de similaridade entre TcExp02B01 e ortólogos de outros eucariotos.

O domínio RAS está sendo demarcado abaixo das sequências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

As proteínas com domínio Rab são encontradas na porção citosólica de membranas intracelulares, sendo que essa localização é variável e determinada por modificações pós-traducionais, em um motivo formado por cisteínas na região carboxi-terminal. Acredita-se que a função desta família de proteínas esteja relacionada com a regulação de vias intracelulares de tráfego vesicular (ZERIAL & McBRIDE, 2001; LANZETTI *et al*, 2004). Mais especificamente, a proteína RAB5 é uma pequena GTPase que está envolvida no controle do tráfego intracelular, atuando como receptor da endocitose e na dinâmica endossomal. Dentro desta perspectiva, Rab5 regula o transporte de vesículas, mediado por clatrina, entre a membrana plasmática e os endossomos primários (recentes), além das fusões entre os endossomos primários (ZERIAL & McBRIDE, 2001).

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *western blot* de TcExp02B01 durante a metaciclogênese pode ser visto na Figura 5.30, onde pode ser observado um aumento, em relação a epimastigotas, de 7,7 vezes em aderido 24 horas e de 1,5 vez em metacíclico. Esse resultado necessita de comprovação devido ao fato de que a intensidade da banda de epimastigota e metacíclico é muito próxima ao ruído de fundo. Embora a quantificação seja menos confiável, o padrão de aumento de expressão em aderido 24 horas é evidente.

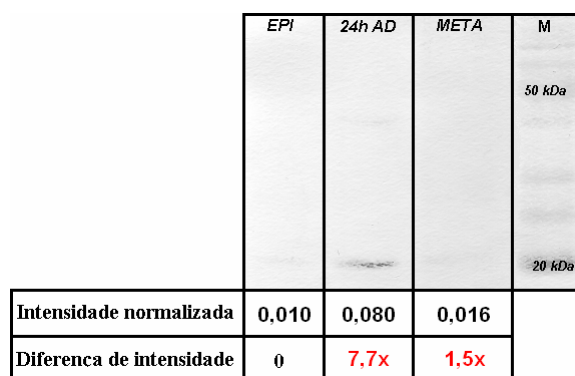


Figura 5.30. Análise de expressão da proteína TcExp02B01 por *western blot* em extratos de *T. cruzi*.
EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho representam aumento da expressão em relação a epimastigotas

Na Figura 5.31, pode-se verificar a localização da proteína TcExp02B01 entre o início do flagelo e o cinetoplasto, em uma região bem definida. Esse local pode ser relativo a estruturas no complexo de Golgi e bolsa paraflagelar.

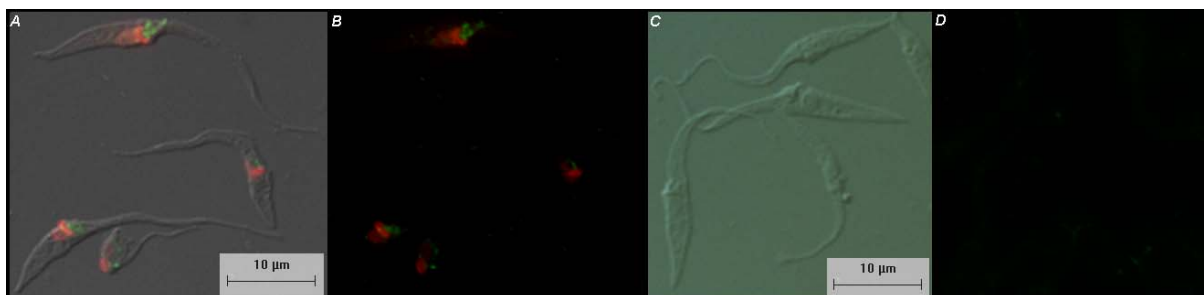


Figura 5.31 Imunolocalização da proteína TcExp02B01.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e marcação de DNA (DAPI, vermelho); C) contraste diferencial do soro pré-imune. D) Imunofluorescência do soro pré-imune.

5.10.9 TcExp02A03, 7109.t00005, Proteína hipotética conservada

Este gene está classificado como o principal do grupo 3 e possui em *T. cruzi* 3027 pares de bases codificando uma proteína hipotética conservada de 113,7 kDa, a qual apresenta um domínio PFAM denominado genericamente de FKBP-type (peptidyl-prolyl Cis-trans isomerase). Proteínas com esse domínio funcionam como receptores de dois imunossupressores, FK506 e rapamicina, e também auxiliam no processo de dobramento (*folding*) de proteínas, catalisando a isomerização cis-trans da prolina (BANG *et al* 2000).

O grau de confiabilidade da determinação desse domínio em *T. cruzi* é relativamente baixo (0,044). Os ortólogos em *T. brucei* e *L. major* apresentam a identificação desse domínio, também com baixa probabilidade. Porém, a sua identificação nos três organismos aumenta a verossimilhança de sua existência, sendo possivelmente degenerado. O início do domínio na sequência gênica relativa ao *T. cruzi* ocorre na posição 129 e termina na posição 232 como mostra a Figura 5.32.

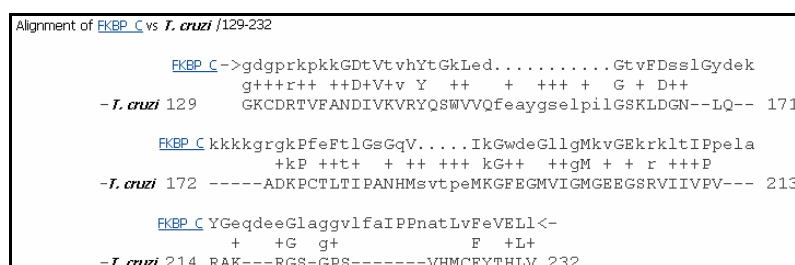


Figura 5.32. Comparação entre TcExp02A03 e o domínioFKBP_C.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 129 da sequência do *T. cruzi* e o término na posição 232.

A sequência deste gene possui homologia com *T. brucei* e *L. major*, de acordo com as sequências obtidas no GeneDB (<http://www.genedb.org>). A Figura 5.33 mostra o alinhamento dessas proteínas demonstrando a similaridade entre os mesmos. Ao se comparar a sequência

de *T. cruzi* com as proteínas presentes no Genbank, foi possível identificar similaridade com proteínas de diversos outros organismos, mas com um grau de conservação mais baixo, impossibilitando a diferenciação entre homologia ou conservação de domínio funcional.

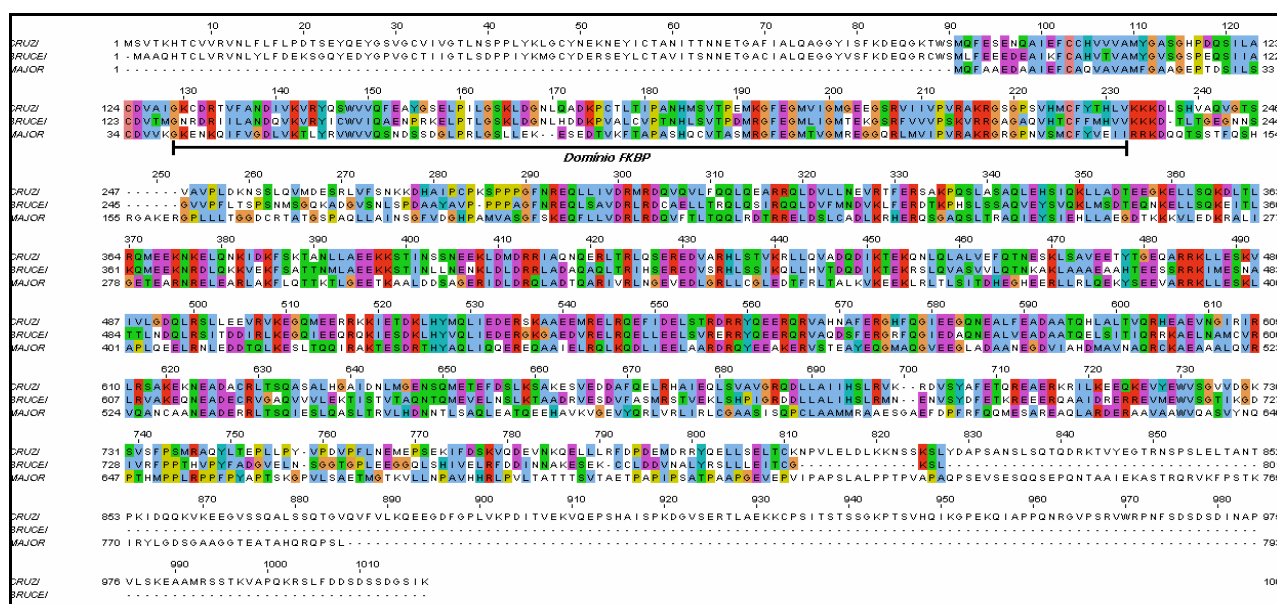


Figura 5.33 Análise de similaridade entre TcExp02A03 e ortólogos de outros tripanossomatídeos. O Domínio FKBP está delimitado. Na parte superior dos alinhamentos encontram-se a numeração relativa à posição aminoacídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém na expressão não obtivemos resultados com os parâmetros já citados no item 4.9.

5.10.10 TcExp02B02, 8764.t00016, Proteína Hipotética

Este gene está classificado como acessório do grupo 3 e possui em *T. cruzi* 1089 pares de bases codificando uma proteína hipotética de 40,3 kDa, apresentando dois domínios denominados de ACP (Phosphopantetheine attachment site) e MYB-2 (*Myb DNA-binding domain repeat signature 2*), preditos pelo programa PROSITE. No entanto, ambos domínios apresentam uma porcentagem de predição de falso-positivos muito alta e, portanto, a determinação dos mesmos não é confiável. Utilizando os programas PFAM e Interpro, não foi possível encontrar domínios conhecidos a partir da sequência de *T. cruzi*.

Inicialmente, esse gene foi classificado como codificando uma proteína hipotética, isto é, sem ortólogos identificados até o momento. No entanto, com a liberação contínua de novos genomas, foi possível identificar uma sequência similar em *L. infantum*, com um *e-value* de $1e10^{-7}$. O alinhamento dessas duas sequências pode ser visto na figura Figura 5.34, evidenciando-se o baixo grau de similaridade entre essas duas sequências, distribuída por

resíduos de aminoácidos dispersos por toda região codificadora, com somente um trecho contendo mais do que três aminoácidos. Porém, a determinação de ortologia é relativamente confiável, pois baseou-se no procedimento de *reciprocal best hit*. Além disso, foi identificado um EST similar a esse gene, o que reforça a sua expressão e verossimilhança como proteína real. Outro dado importante para reforçar esse resultado é a sua localização em um fragmento cromossômico sem indícios de erro de montagem, o que muitas vezes cria artificialmente proteínas inexistentes, as quais são classificadas como hipotéticas.

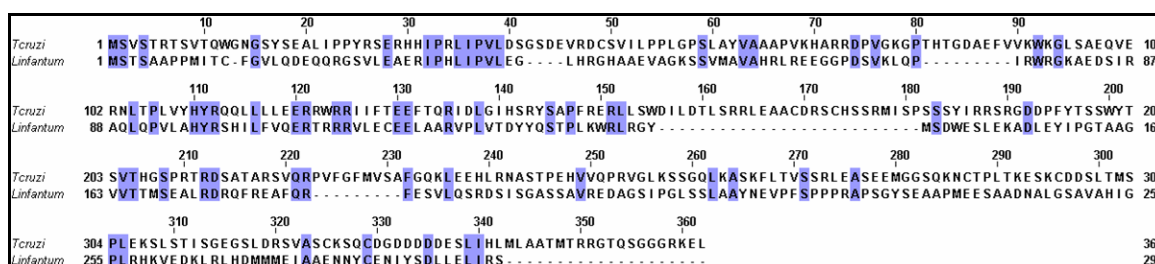


Figura 5.34 Análise de similaridade entre seqüências de *T. cruzi* e *L. infantum*.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão. A expressão foi obtida, bem como o soro policlonal, porém só houve reconhecimento da proteína recombinante. Quando utilizado extrato de *T. cruzi* não apresentou sinal.

5.10.11 TcExp02B03, 8728.t00021, Proteína Hipotética Conservada

Este gene está classificado como secundário do grupo 3 e possui em *T. cruzi* 954 pares de bases codificando uma proteína hipotética conservada de 35,1 kDa, apresentando um domínio PFAM denominado zf-C2H2 (*Zinc Finger Protein C2H2*). Este domínio representa o *zinc finger* clássico e possui função ligadora a ácidos nucleicos (DNA e RNA) fazendo parte de muitas proteínas associadas à expressão gênica. A característica principal deste domínio é a conservação de duas cisteínas e duas histidinas. Na Figura 5.35 pode-se verificar a conservação das cisteínas nas posições 3 e 7 e da histidina nas posições 20 e 24 na representação em logo-plot do modelo do domínio disponível na base de dados PFAM.

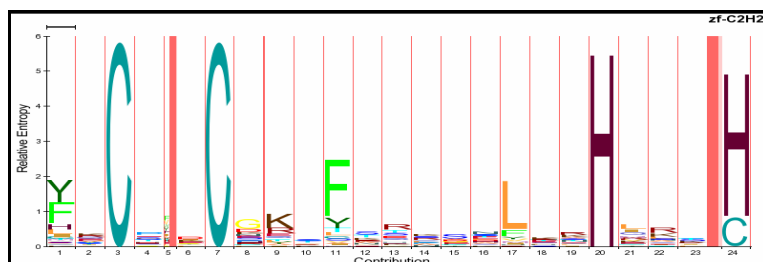


Figura 5.35 Gráfico HMM relativo o domínio zf-C2H2.

Gráfico HMM retirado do programa PFAM mostrando o grau de conservação em cada posição aminoacídica do domínio zf-C2H2. No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

A comparação entre o domínio zf-C2H2 e TcExp02B03 pode ser vista na Figura 5.36, e o valor de confiabilidade foi $1,3e^{-5}$. O início do domínio na sequência gênica relativa ao *T. cruzi* ocorre na posição 79 e termina na posição 102, evidenciando-se a conservação das duas cisteínas e as duas histidinas.

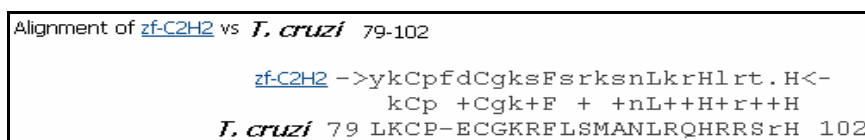


Figura 5.36 Comparação entre a sequência de *T. cruzi* e o domínio zf-C2H2.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 71 da sequência do *T. cruzi* e o término na posição 102.

A comparação da sequência TcExp02B03 com o proteoma de outros organismos identificou ortólogos apenas em tripanossomatídeos. O alinhamento desses ortólogos obtido pelo programa Clustal W é mostrado na Figura 5.37, podendo-se verificar a conservação das duas cisteínas bem com das duas histidinas características do domínio em questão. Além do domínio, este gene apresenta conservação nas outras regiões em relação às outras espécies, sendo que as proteínas de *Leishmania* apresentam um tamanho maior.

Alignment of DREV vs *T. cruzi* [74-344]

```

DREV->WYvlnreKLseslQavFvaslpDaeTqsFLenskklsanlFiqlzha
      Y +n L+ + + v D eT +L +S +a
-T. cruzi 74 -YTPNLVCLPNHIVSKHVSVMCMDBETTEWLRAGDVS-----VMKA 113

DREV LassiLsLfmskTdINGllgrGsMFvfSeeQFrkLLaikespassdqkd
L s iL f+s+T NG+l rG MFv+S Q + LL + ss +
-T. cruzi 114 LLSGILRAFFSQTTSNGILNRGGMFVMSTAQAQALLGNEDDQRSSKSP 163

DREV k.kLLDLGAGdGeitkrmapffeeLSGSAGLRVYATElSetMrdrLkkkk
++lLD+GAGdG +t r+ p f + V TE S Mr+rL ++
-T. cruzi 164 FgTLLDIGAGDGGVTSRLQPLFQH-----VAVTEYSIPMRWRLWQR 205

DREV YnVLSaleGwqeTd..vkNydlilcLNlLDRcydplkLLeDirsvalept
Y VL + + d++ +yd+i cLN+LDR + pl LL +r l p
-T. cruzi 206 YEVLFPYQDPFRHKDGSRRYYDVIACLNVLDRADKPELLRSMRD-SLSP- 253

DREV nGrvivALVLPyshYVEtntnGksekpdleLvkG...rqrseyeevar
nG +A VLP+ ++VE++ + +p e L G++ r++syet r
-T. cruzi 254 NGILLAVVLPWCPFVNGI--RQRRTETLPMEGGecCGASyEDSLQR 301

DREV lmE..llrkaGyvveAwTlPYLCEGdmynsfYyLdDaifvLkp<-
l ++l +G++v w +lPY CEG + + +L Da+fvL
-T. cruzi 302 LVDnVLI PCGFKVDRWCKLPYFCENLRVEYAVLHDAVFVLT 344

```

Figura 5.38 Comparação entre a sequência de *T. cruzi* e o domínio DREV.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 74 da sequência do *T. cruzi* e o término na posição 344.

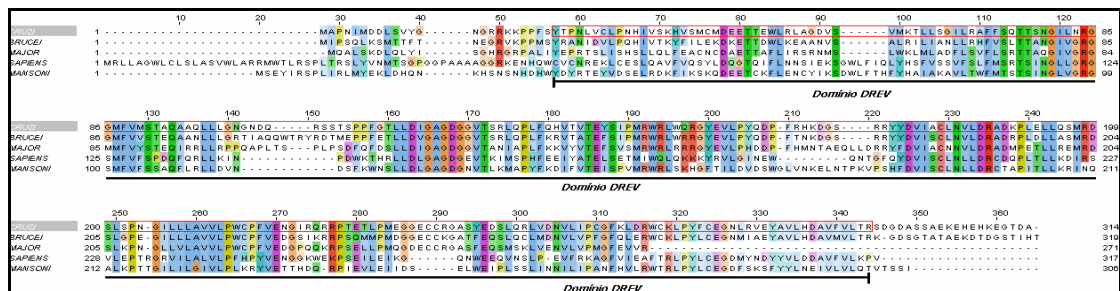


Figura 5.39 Análise de similaridade entre TcExp02B04 e os ortólogos de outros eucariotos.

O Domínio DREV está sendo demarcado abaixo das sequências analisadas. Na parte superior dos alinhamentos encontram-se a numeração relativa à posição nucleotídica.

Este gene foi clonado, obtendo-se o vetor de entrada e de expressão. A proteína foi expressa, porém a obtenção do soro não foi possível devido à morte dos animais durante as inoculações, não sendo possível a realização dos ensaios de *western blot* e nem imunolocalização por microscopia óptica.

5.10.13 TcExp02A04, 6896.t00027, Proteína Hipotética conservada

Este gene está classificado como principal do grupo 4 e possui em *T. cruzi* 1353 pares de bases codificando uma proteína hipotética conservada de 51,4 kDa, apresentando um domínio denominado SWIM (*Zinc Finger Protein* SWIM, e-value de $3,2 \times 10^{-3}$) e outro como zf-C3HC4 (*Zinc Finger Protein* C3HC4 ou *RING finger*, e-value de $7,9 \times 10^{-4}$), ambos preditos a partir do banco de dados PFAM. Este último domínio não foi identificado pelo consórcio de anotação do genoma de *T. cruzi*. O primeiro domínio citado é encontrado em bactérias, arqueobactérias e eucariotos, e sua conformação pode ser semelhante a dos domínios *Zinc*

finger zf-C2H2, sendo um domínio versátil, podendo interagir com DNA ou proteínas em diferentes contextos (MAKAROVA *et al.*, 2002). Já o domínio RING *finger* está presente em proteínas que participam da via de ubiquitinação. A comparação dos domínios identificados com a sequência de TcExp02A04 pode ser observada na Figura 5.40. O domínio SWIM está localizado entre os resíduos 49 e 79, e o zf-C3HC4 entre os resíduos 315 e 356.

Esse gene apresenta alguns outros domínios, parciais e com baixo valor de confiabilidade (degenerados ou preditos erroneamente), que são funcionalmente relacionados aos domínios supracitados: PHD (resíduos 131 a 181, e-value de 0,099), zf-B Box (resíduos 202 a 234, e-value de 0,68). Isso pode indicar uma maior especialização dessa proteína, porém com modificações dos domínios acessórios tornando sua identificação problemática.

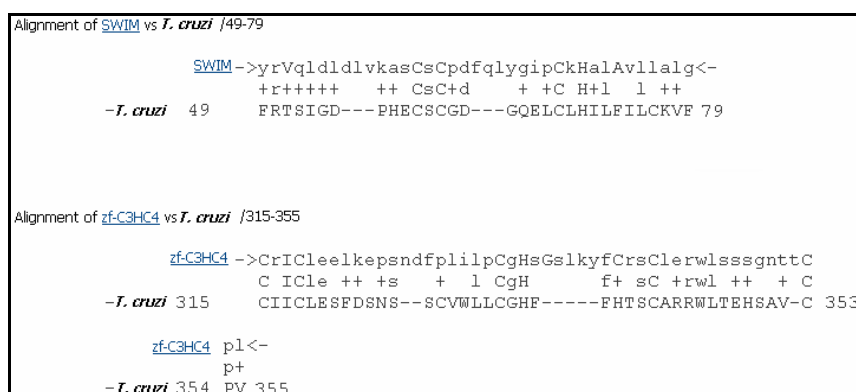


Figura 5.40 Comparação entre a sequência de *T. cruzi* e os domínios SWIM e zf-C3HC4 utilizando o programa PFAM.

Entre as duas sequências encontra-se a conservação aminoacídica.

O domínio SWIM possui como característica a conservação de 3 cisteínas e uma lisina como pode ser visto na Figura 5.41. Já o domínio C3HC4 (RING) possui 3 cisteínas conservadas seguida de uma lisina e 4 cisteínas, (Figura 5.42). Nota-se a grande variação nas posições aminoacídicas das regiões adjacentes às posições conservadas em ambos os domínios.

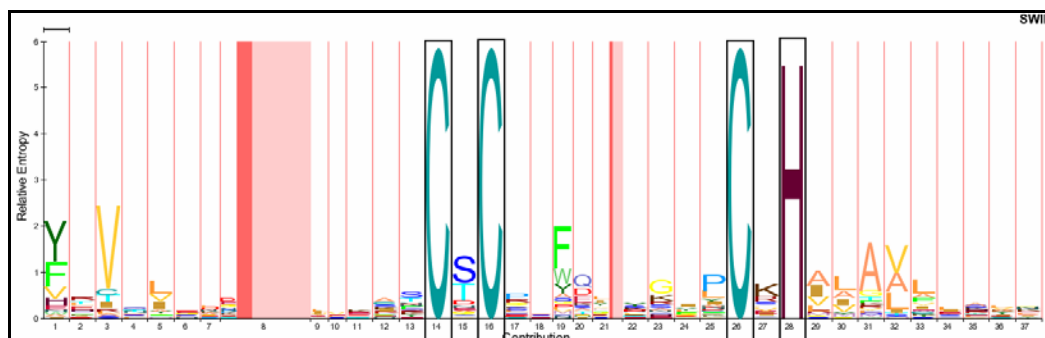


Figura 5.41. Gráfico HMM relativo ao domínio SWIM

Gráfico HMM retirado do programa PFAM mostrando o grau de conservação em cada posição aminoacídica do domínio SWIM. No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa. Os aminoácidos C (Cisteína) e K (Lisina) característicos do domínio estão demarcados por retângulos pretos.

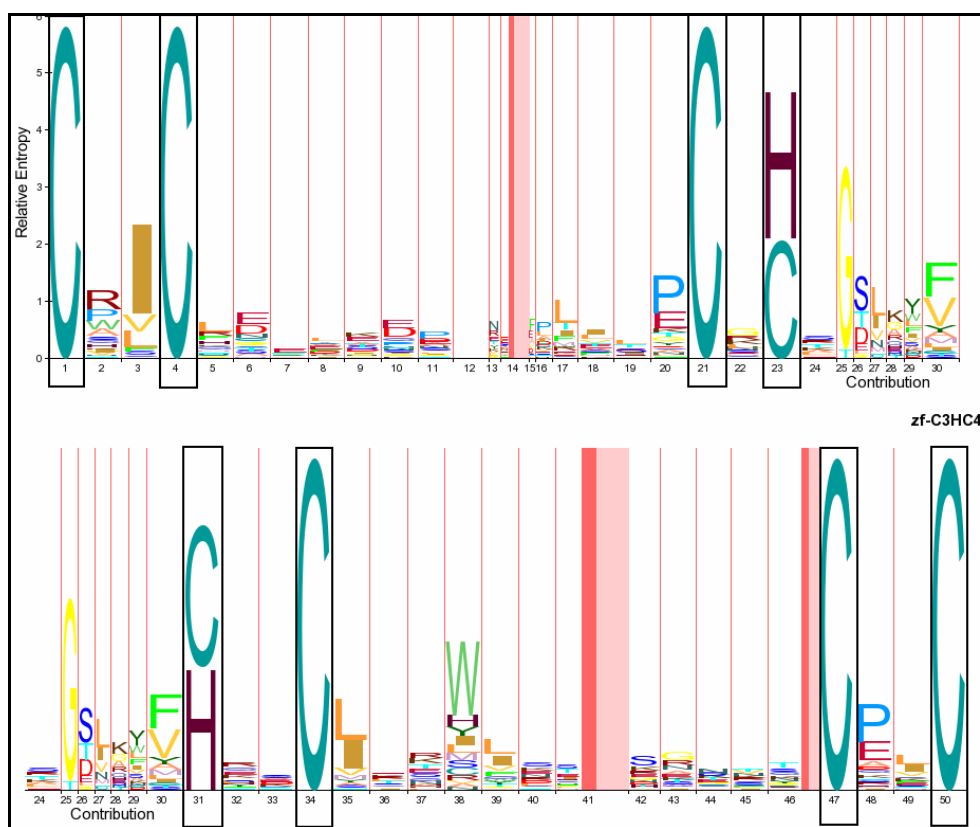


Figura 5.42. Gráfico HMM do domínio zf-C3HC4 (RING).

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa. Os aminoácidos C (Cisteína) e K (Lisina) característicos do domínio estão demarcados por retângulos pretos.

A comparação da sequência de *T. cruzi* com as proteínas presentes no Genbank, utilizando o programa BLASTp, identificou várias sequências com um valor de significância próximo a $1e^{-28}$, sendo selecionadas duas sequências que aparecem também no programa Homologene. Esse domínio é conservado desde leveduras até humanos. A análise de

seqüências homólogas pode ser vista na Figura 5.43, podendo se verificar um padrão de conservação dos dois domínios, bem como outras regiões de conservação.

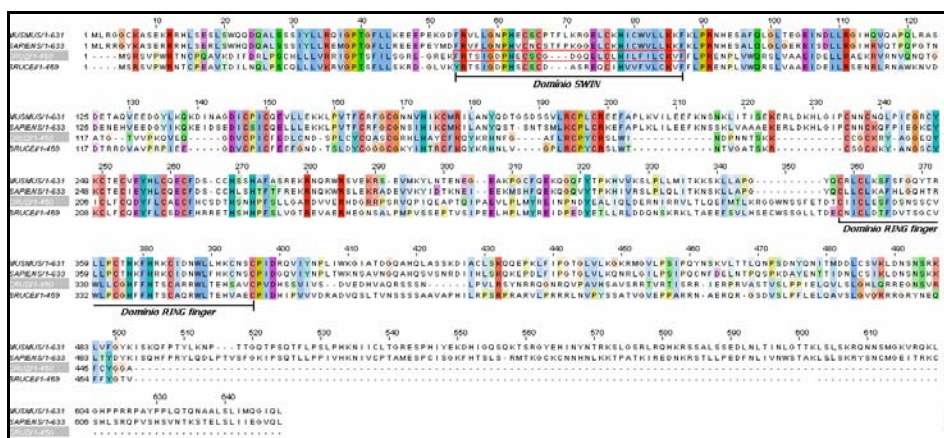


Figura 5.43 Análise de similaridade entre TcExp02A04 e ortólogos de outros eucariotos. O Domínio SWIM e RING finger estão sendo demarcados abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.44, onde pode ser observado um forte aumento, em relação a epimastigotas, de 4,5 vezes em aderido 24 horas e uma diminuição de 3,1 vezes em metacíclicos. A magnitude desse último resultado deve ser considerada com cautela devido ao sinal fraco de epimastigota e metacíclico. No entanto, o padrão de aumento em 24 horas e diminuição em metacíclicos é claro.

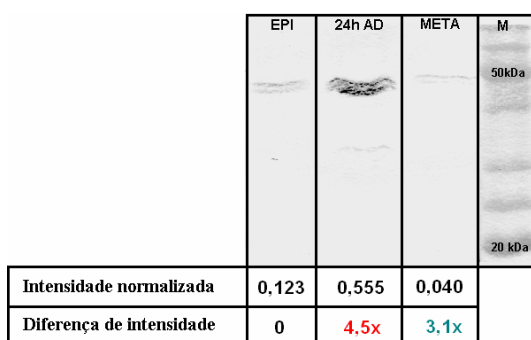


Figura 5.44. Análise de expressão da proteína TcExp02A04 por *western blot* em extratos de *T. cruzi*. EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Valores em vermelho e verde representam, respectivamente, aumento e diminuição em relação a epimastigotas

A imunolocalização foi realizada, porém apresentou reação cruzada com o soro pré-imune, sendo assim, descartado o resultado a princípio.

5.10.14 TcExp02B05, 7224.t00006, Proteína Hipotética

Este gene está classificado como acessório no grupo 4 e possui em *T. cruzi* 1485 pares de bases codificando uma proteína hipotética de 52,7 kDa. Este gene não apresentou domínios previstos pelos programas PFAM e Interpro. Não foram identificados ortólogos em outros organismos, mesmo incluindo os genomas recentemente seqüenciados. Esse gene apresenta uma região poli-valina, entre os aminoácidos 45 e 50, a qual difere de tamanho entre os dois alelos de CL Brener (6895.t00004 e 7224.t00006), além de duas regiões ricas em glutamina, nas regiões 306 a 314 e 407 a 422, conforme pode ser visto na Figura 5.45. A implicação funcional dessas seqüências não está estabelecida.

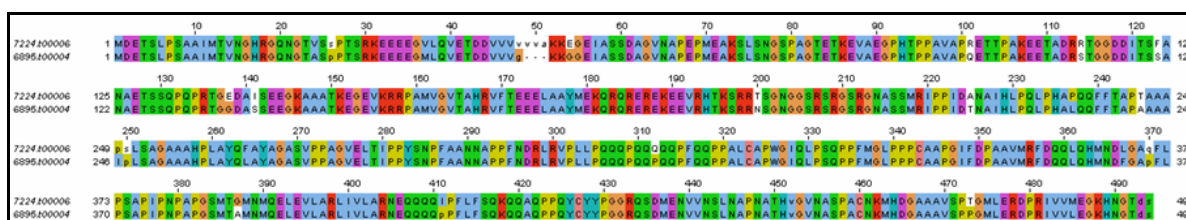


Figura 5.45. Alinhamento das proteínas de TcExp02B05 presentes em CL Brener.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém obtivemos uma expressão fraca de um produto com um tamanho maior do que o esperado, 90 kDa. O anticorpo não reconheceu nenhum sinal no extrato de *T. cruzi*.

5.10.15 TcExp02B06, 6958.t00014, Proteína Hipotética Conservada

Este gene é o secundário do grupo 4 e possui em *T. cruzi* 249 pares de bases codificando uma proteína hipotética de 9,2 kDa. Apresenta um domínio do tipo *Zinc finger* denominado CSL. Este domínio está freqüentemente presente associado a DnaJ e também é encontrado nas proteínas DPH3 e DPH4, as quais estão envolvidas na biossíntese de diftâmida, um resíduo de histidina modificado pós-traducionalmente, encontrado somente no fator de elongação da tradução 2 (eEF-2), sendo conservada de arqueobactérias a humanos e que serve como alvo para a toxina da difteria (LIU *et al*, 2004).

Em *T. cruzi*, apenas esse gene foi identificado com confiabilidade (e-value de $1,3e^{-10}$) como contendo o domínio zf-CSL. Este domínio é caracterizado pela conservação de 4 cisteínas, seguidas de uma leucina e de uma serina, como pode ser observado Figura 5.46. A região de similaridade entre o domínio zf-CSL e a proteína TcExp02B06 pode ser vista na

Figura 5.47, a qual se estende por uma grande porção da proteína (dos resíduos 7 a 59, de um total de 83 aminoácidos).

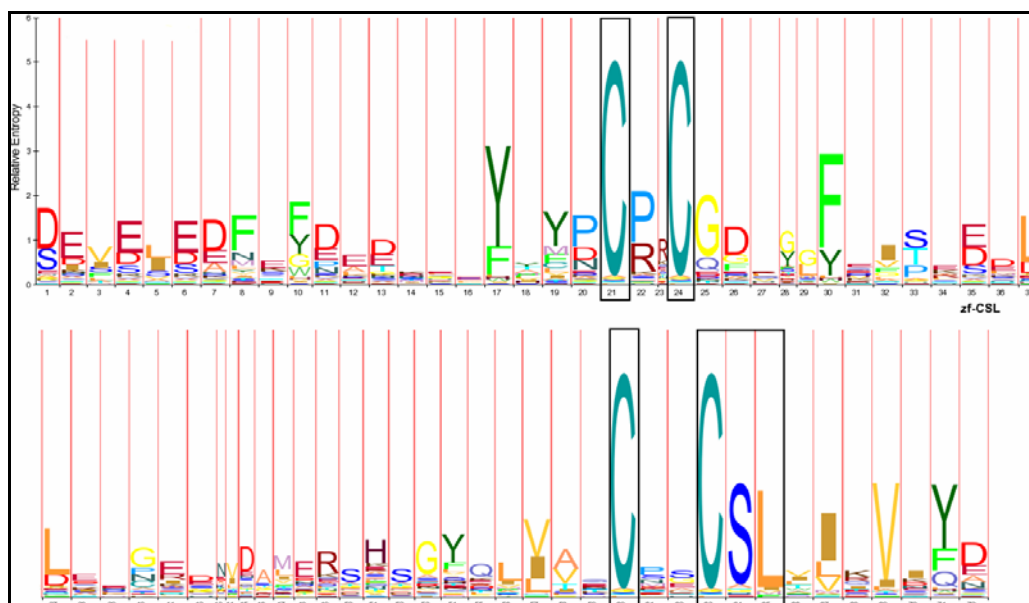


Figura 5.46 Gráfico HMM do domínio zf-CSL.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa. Os aminoácidos com maior valor de conservação característicos do domínio estão demarcados por retângulos pretos.

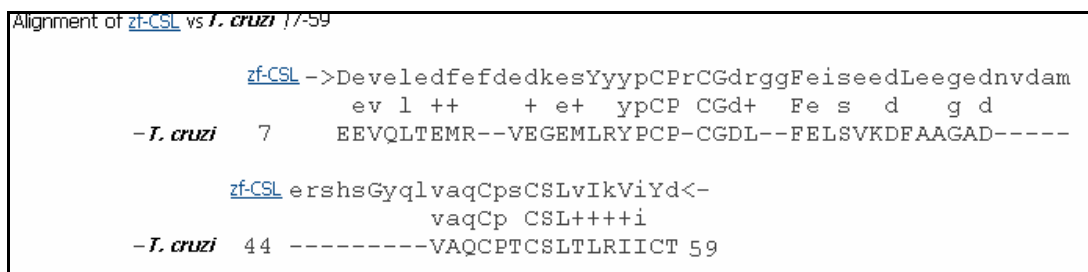


Figura 5.47 Comparação entre a sequência de *T. cruzi* e o domínio zf-CSL.

Entre as duas sequências encontra-se a conservação aminoacídica.

Ao se realizar uma comparação da proteína TcExp02B06 com todas as proteínas presentes no Genbank, através de um BLASTp, foi possível identificar possíveis ortólogos em diversos organismos, como *L. major*, *T. brucei*, *A. thaliana*, *H. sapiens*, *S. cerevisiae* entre outros eucariotos, com valor de significância próximo a $1e-17$ para tripanossomatídeos e $1e-7$ para outros organismos. O alinhamento das seqüências homólogas pode ser visto na Figura 5.48, no qual é possível verificar que a maior similaridade está concentrada na região relativa ao domínio zf-CSL. Todas essas proteínas apresentam um tamanho semelhante e estão anotadas, geralmente, como DPH3 ou KTI11, isto é, uma das proteínas necessárias para a síntese de diftamida, conforme mencionado acima.

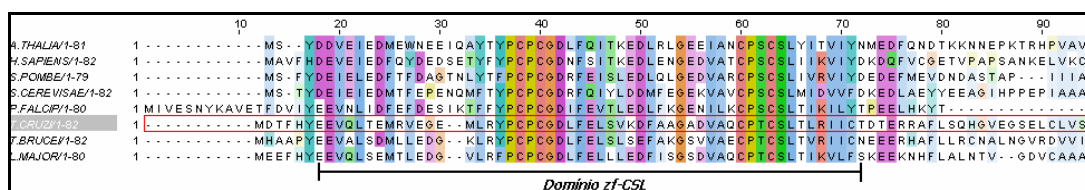


Figura 5.48 Análise de similaridade entre seqüências de *T. cruzi* e outros eucariotos.

O Domínio zf-CSL está sendo demarcados abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se a numeração relativa à posição nucleotídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão. A expressão foi obtida, bem como o soro policlonal, porém com uma qualidade baixa, necessitando do método de quimioluminescência para detectar a proteína no ensaio de *western blot*. Apresentou um aumento de intensidade de sinal na fase tripomastigota metacíclico em relação a forma epimastigota, de 1,5 vez, como pode ser verificado na Figura 5.49.

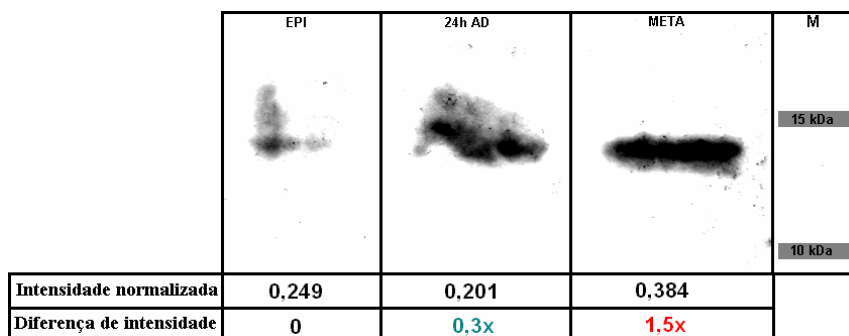


Figura 5.49 Análise de expressão da proteína TcExp02B06 por *western blot* em extratos de *T. cruzi*.

EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho e verde representam, respectivamente, aumento e diminuição da expressão em relação a epimastigotas

A imunolocalização foi realizada, porém a permeabilização das células que ocorre durante o processo de confecção das lâminas não ocorreu de forma adequada, impedindo a obtenção de resultados. Este procedimento será repetido futuramente.

5.10.16 TcExp02B07 8243.t00004, Dynein light chain 2B, cytoplasmic, putative

Este gene está classificado como guia do grupo 4 e possui em *T. cruzi* 354 pares de bases codificando uma proteína hipotética conservada de 13,3 kDa, apresentando um domínio PFAM denominado de Roadblock/LC7, com um valor de confiabilidade de 0,007, que se estende por praticamente toda a proteína, dos resíduos 18 a 113, de um total de 118 aminoácidos (Figura 5.50). Este domínio está presente em uma família de proteínas que fazem parte do complexo multimérico da dineína, mais precisamente encontrada no braço externo do

complexo, podendo estar associada com diferentes proteínas resultando em função reguladora específica da dineína (PFISTER *et al*, 2006).

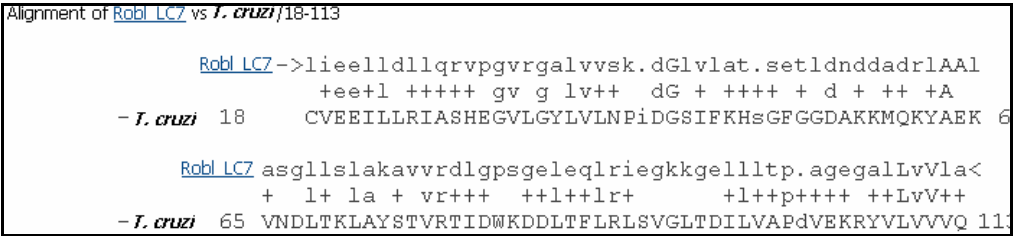


Figura 5.50. Comparação entre a sequência de *T. cruzi* e o domínio Roadblock/LC7. Entre as duas sequências encontra-se a conservação aminoacídica.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências com um valor de significância próximo a $1e^{-35}$ para tripanossomatídeos e $1e^{-9}$ para outros organismos, sendo que esse domínio é conservado desde leveduras até humanos. A comparação das proteínas homólogas pode ser vista na Figura 5.51.

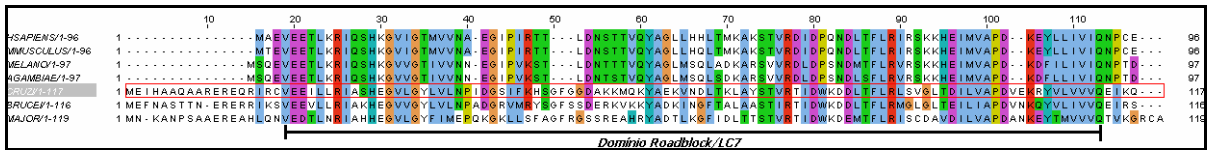


Figura 5.51. Análise de similaridade entre TcExp02B07 e ortólogos de outros eucariotos. O Domínio Roadblock/LC7 está sendo demarcados abaixo das sequências analisadas. Na parte superior dos alinhamentos encontram-se a numeração relativa à posição nucleotídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *Western blot* da metaciclogênese pode ser visto na Figura 5.52, onde pode ser observada um aumento de duas vezes em aderido 24 horas e de 4,3 vezes em metacíclicos, em relação a epimastigotas.

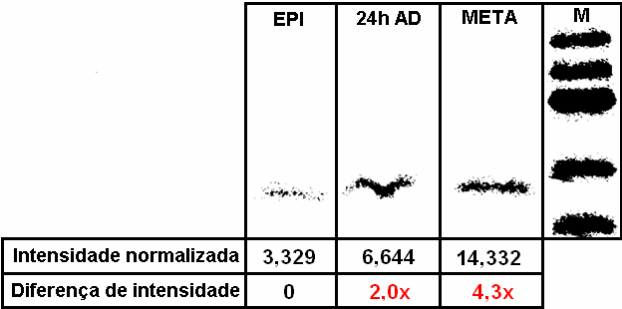


Figura 5.52. Análise de expressão da proteína TcExp02B06 por *western blot* em extratos de *T. cruzi*. EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho representam aumento da expressão em relação a epimastigotas

O padrão de localização encontrado para a proteína TcExp02B07, pode ser visto na Figura 5.53, onde verifica-se a existência de fluorescência relativamente baixa no citoplasma, com grânulos mais intensos na região do flagelo.

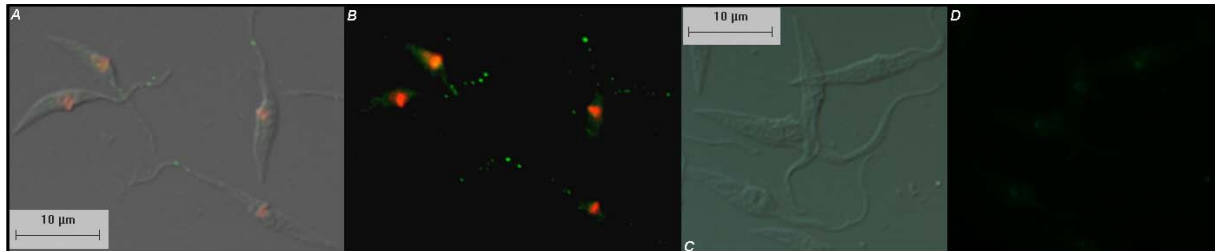


Figura 5.53 Imunolocalização de TcExp02B07.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e marcação do DNA (DAPI, vermelho); C) contraste diferencial do soro pré-imune. D) Imunofluorescência do soro pré-imune.

5.10.17 TcExp02A05, 4715.t00001, Anti-silencing protein, ASF1-like

Este gene está classificado como o principal do grupo 5 e possui em *T. cruzi* 606 pares de bases codificando uma proteína hipotética conservada de 22,2 kDa, apresentando um domínio PFAM denominado de ASF1-LIKE (e-value de $4,9e^{-13}$), o qual constitui grande parte da proteína, se estendendo do resíduo 1 até o 150, de um total de 202 aminoácidos (Figura 5.54). Este domínio é bastante conservado evolutivamente e está relacionado com a inibição de repressão gênica, atuando como chaperona de histonas, interagindo com as histonas H3 e H4 e o fator de deposição de histonas CAF-I. (MUNAKATA *et al*, 2000; DAGANZO *et al*, 2003). O gene CAF-I está presente em *T. cruzi* (4574.t00010/5857.t00006).

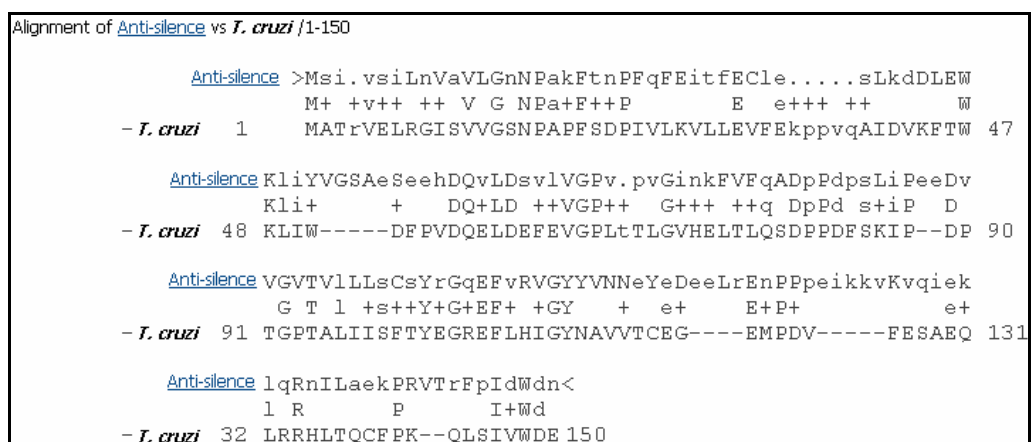


Figura 5.54. Comparação entre a sequência de *T. cruzi* e os domínios Anti-silence. Entre as duas sequências encontra-se a conservação aminoácídica.

O domínio Anti-Silence apresenta uma conservação de vários aminoácidos como pode-se verificar na Figura 5.55, como: prolina, fenilalanina, tirosina, glicina, arginina, ácido glutâmico e alanina.

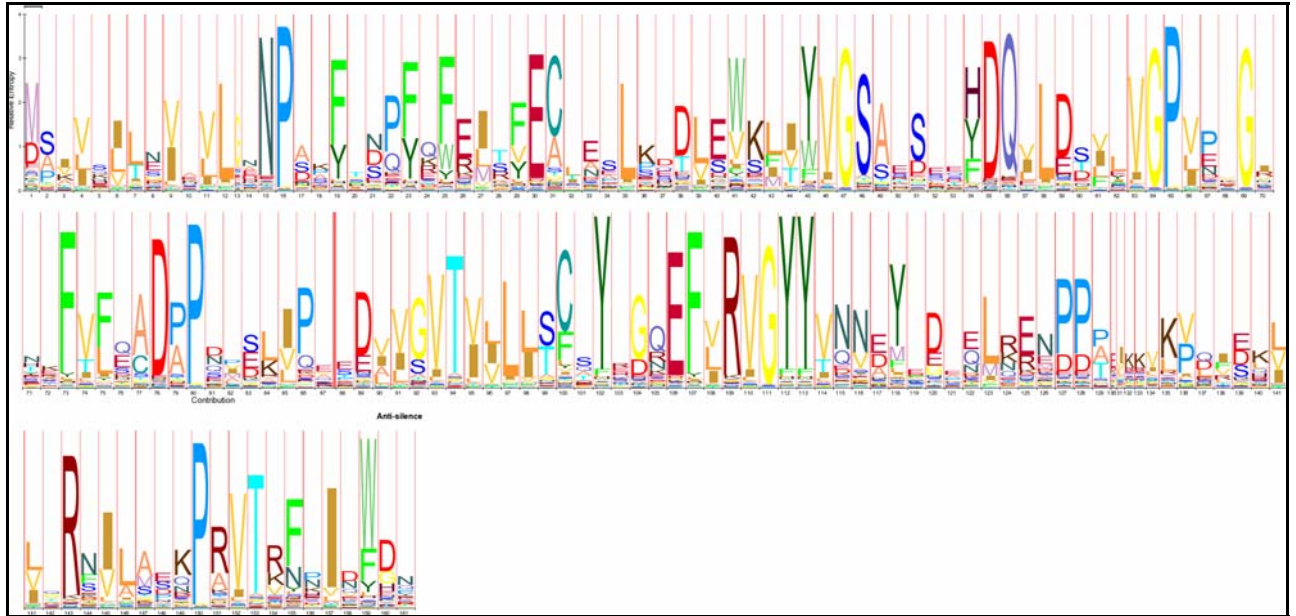


Figura 5.55. Gráfico HMM do domínio Anti-silence.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências com um valor de significância próximo a $1e^{-55}$ para tripanossomatídeos e $1e^{-9}$ para outros organismos selecionados, sendo que esse domínio é conservado de leveduras a humanos. A comparação das proteínas homólogas pode ser vista na Figura 5.56.

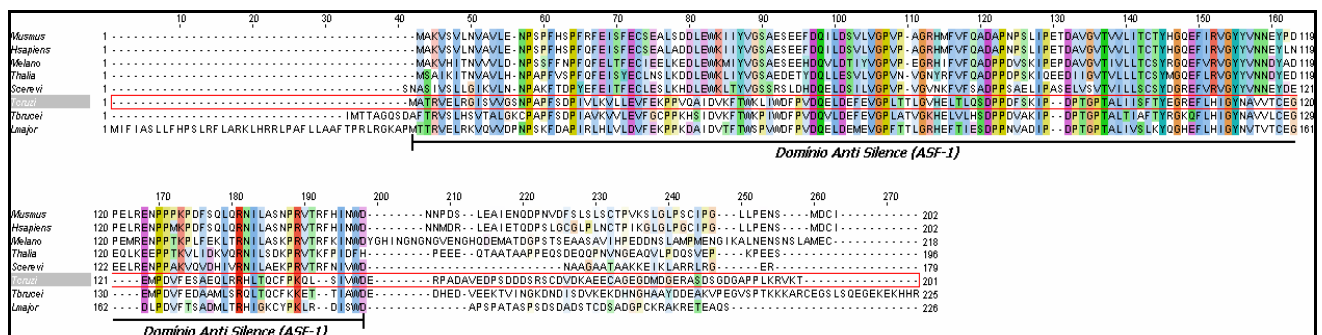


Figura 5.56. Análise de similaridade entre TcExp02A05 e ortólogos de outros eucariotos.

O Domínio Anti-silence está sendo demarcados abaixo das sequências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão. A expressão foi obtida, bem como o soro policlonal, porém só houve reconhecimento da proteína recombinante. Quando utilizado extrato de *T. cruzi* não apresentou sinal adequado.

5.10.18 TcExp02B08, 5595.t00003, Proteína hipotética

Este gene está classificado como acessório do grupo 5 e possui em *T. cruzi* 516 pares de bases codificando uma proteína hipotética de 19,5 kDa, não apresentando domínios conhecidos nos bancos de dados PFAM, Interpro e Prosite. Apresenta a predição de quatro regiões transmembranas que se estendem por praticamente toda a proteína; sendo esta composta por aminoácidos hidrofóbicos em mais de 50% de seus resíduos, tendo 14 resíduos básicos que são encontrados geralmente em diaminoácidos, conforme visto na Figura 5.57.

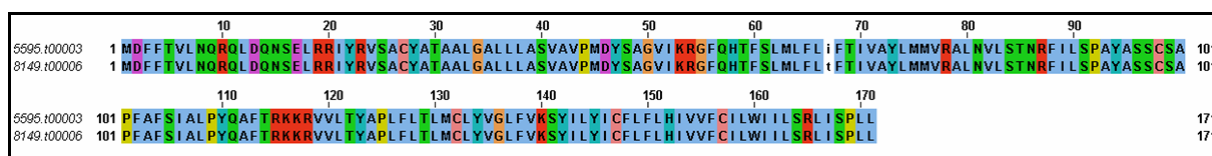


Figura 5.57. Análise de similaridade entre duas seqüências de *T. cruzi* CL Brener.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém não foi obtida a expressão dessa proteína.

A princípio, esses resultados indicariam uma possível falha na predição desse gene, pois é uma proteína somente identificada em *T. cruzi*, apresenta uma hidrofobicidade grande e não foi obtida expressão dessa proteína em *E. coli*. No entanto, ao se verificar a sonda que hibridou com esse gene, foi identificado um EST (código de acesso AI717787) que se localiza no final da região codificadora e no início da região 3'UTR do mRNA, sugerindo a sua expressão natural. Além disso, com a liberação da seqüência do genoma de outros tripanossomatídeos, foi possível identificar possíveis ortólogos em *T. congolense* e *T. vivax*, conforme pode ser visto na Figura 5.58. Não foram identificados ortólogos nos outros tripanossomatídeos: *T. brucei*, *T. gambiense*, *L. major*, *L. infantum* e *L. brasiliensis*. Esses dados levam a mudar a anotação dessa proteína para hipotética conservada.

Esses resultados reforçam a noção de que essa proteína é possivelmente funcional, estando conservada em alguns tripanossomatídeos, tendo sido perdida nos demais. A ausência de expressão em *E. coli* pode ser devido a seu caráter hidrofóbico, o que reforça a importância do estabelecimento de sistemas de expressão mais próximos evolutivamente a *T. cruzi*.

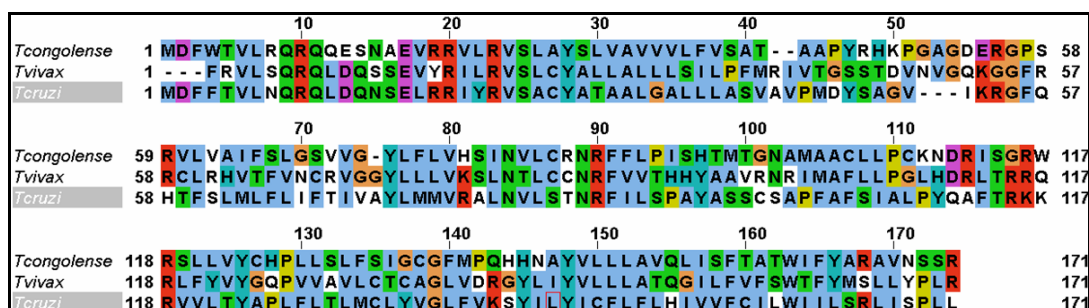


Figura 5.58. Análise de similaridade entre seqüências de *T. cruzi*, *T. congolense* e *T. vivax*.

5.10.19 TcExp02B09, 8171.t00006, Proteína Hipotética conservada

Este gene está classificado como secundário do grupo 5 e possui em *T. cruzi* 1035 pares de bases codificando uma proteína hipotética conservada de 38,2 kDa, apresentando um domínio PFAM denominado de PDC2-C (Programmed cell death protein 2, C-terminal). A proteína PDCD2 está localizada no citosol de células situadas no pólo oposto do centro germinativo de centoblastos; e o homólogo em rato (RP8) foi relacionado com a morte celular programada em timócitos. Foi demonstrado que interage com BCL6, um fator transcricional do tipo *Kruppel zinc finger*, que é conservado evolutivamente. No entanto, não foi identificado ortólogo de BCL6 em *T. cruzi*, tanto usando a similaridade primária de seqüência pelo BLASTp quanto pela procura dos domínios presentes em BCL6 nas proteínas de *T. cruzi* (domínios zf-C2H2 e BTB), os quais estão presentes em algumas proteínas de *T. cruzi*, mas não são encontrados juntos.

O domínio PDC2-C apresenta uma conservação de vários aminoácidos como pode ser verificado na Figura 5.59.

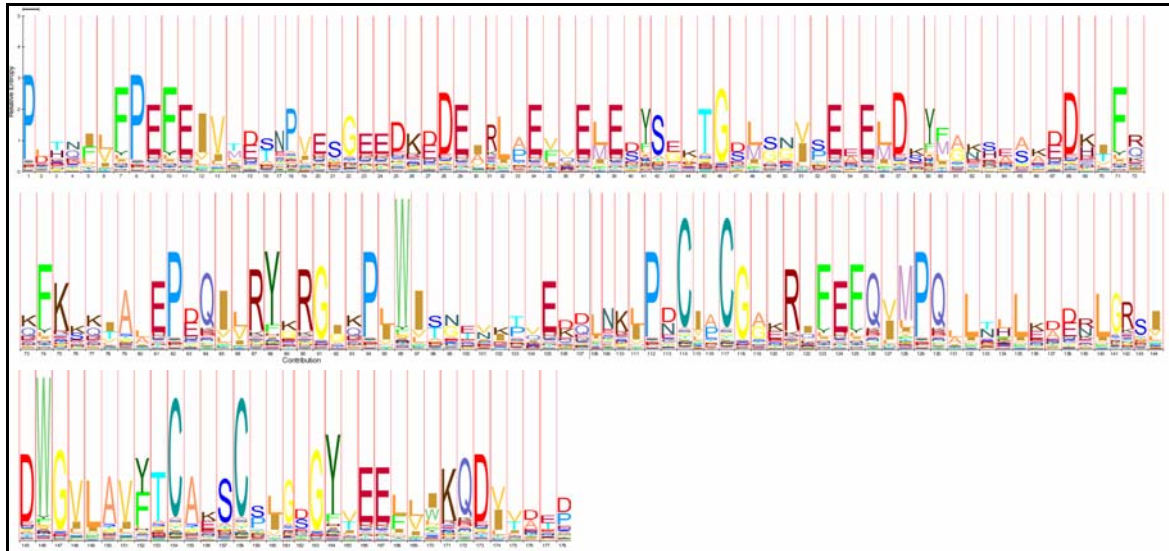


Figura 5.59. Gráfico HMM do domínio PDC2-C.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

Quando comparada a similaridade entre o domínio PDC2_C e a sequência do *T. cruzi*, verificou-se um valor de confiabilidade igual a $6e^{-7}$. O início do domínio na sequência proteica de *T. cruzi* ocorre na posição 131 e termina na posição 319 como mostra a Figura 5.60.

PDCD2_C vs <i>T. cruzi</i> 131-319	
PDCD2_C	->PltnilFPFEFEIVtdsnPvEsGEEDKdDEiRLa.Ev.vEEdYSeKT
	P + + EIV + Ei ++E++ E+ Se
<i>T. cruzi</i> 131	EPHTFPFCFVEIVAEP-----KKEIVVPtELEAEMIRASEEK 167
PDCD2_C	GdlsnvsEe...EldKyfakseakdDkiFrkFKkkiALEPdQILRY...
	++ E++ +El + + + D+ F kF+++ A EP Q+LRY ++
<i>T. cruzi</i> 168	ARNAGITEAdikeLERTIDLKDKPSDHEFEKFRRRMAREPRQVLRyvrD 217
PDCD2_CkRG.....ikPlWIsnevktvEddLnKlPdCiaCGakRiFEFQvM
	+ +++k G + ++ Pl + +P C CGa E Q M
<i>T. cruzi</i> 218	plgkkKNGapaptASPLFMHPRNVGNRAR---IPPCDSCGAALMHELQLM 264
PDCD2_C	PQlLtllLkddrL...GRs.....iDWGVLAVyTCAkSCslGdGYvEelli
	P +L+ +G+s +++DWG V+ C k Cs ++
<i>T. cruzi</i> 265	PTCVYYLRVGEYiatGKSaadegVDWGSVTVFVCSKDCSK---DCSGVVL 311
PDCD2_C	KQDvvded<-
	+ v +
<i>T. cruzi</i> 312	RRAFVLVE 319

Figura 5.60. Comparação entre a sequência de *T. cruzi* e o domínio PDC2_C.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 131 da sequência do *T. cruzi* e o término na posição 319.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou ortólogos possíveis, geralmente anotados como *programmed cell death 2*, em outros eucariotos (Figura 5.61).

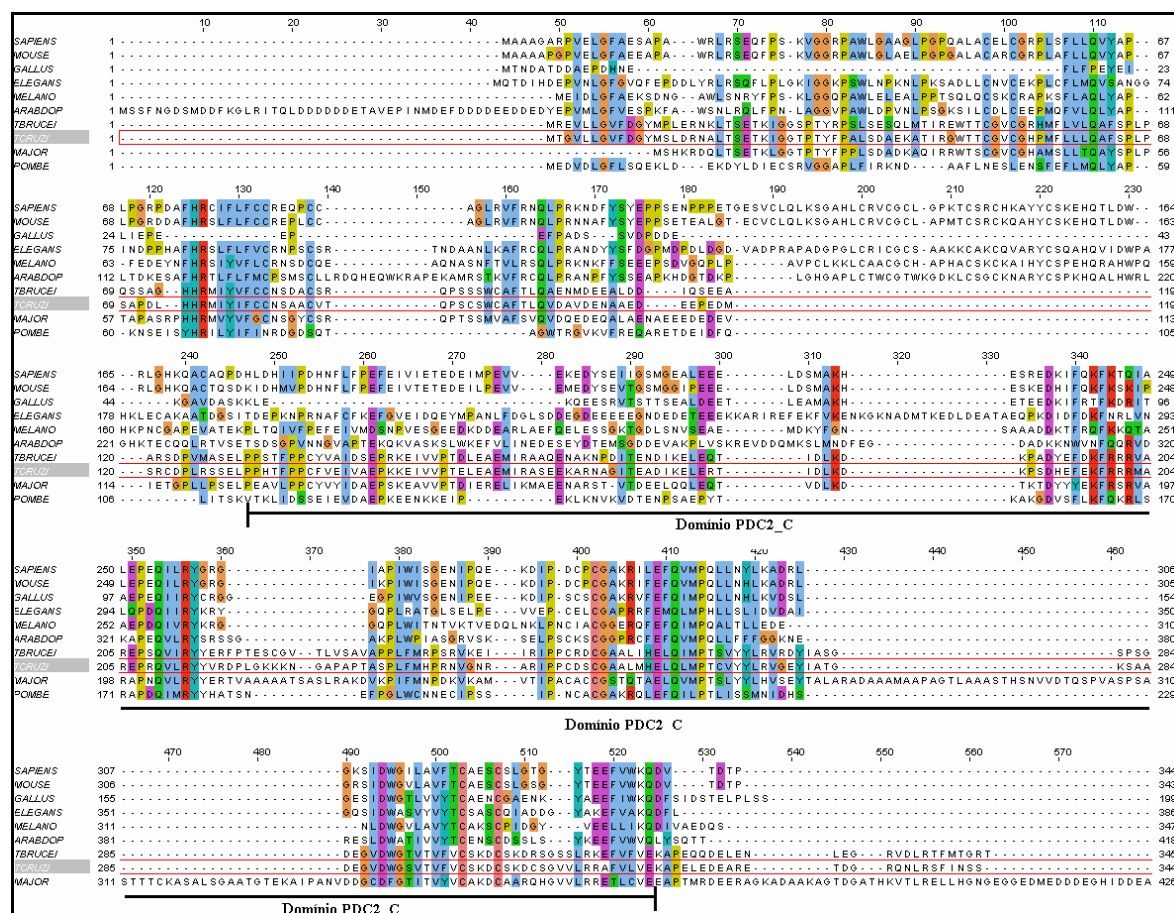


Figura 5.61. Análise de similaridade entre TcExp02B09 e ortólogos de outros eucariotos.
O Domínio PDC2_C está sendo demarcado abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão. A expressão da proteína foi obtida, bem como o soro policlonal. No entanto, só houve reconhecimento da proteína recombinante, pois quando utilizado extrato de *T. cruzi* não apresentou sinal.

5.10.20 TcExp02B10, 8416.t00006, Calcineurin B subunit, putative

Este gene está classificado como guia do grupo 5 e possui em *T. cruzi* 531 pares de bases codificando uma proteína de 19,5 kDa, apresentando três domínios denominados EF-hand. Esse domínio pequeno, de 29 aminoácidos (Figura 5.62), está presente em um extenso grupo de proteínas que possuem a função geral de ligação ao cálcio, as quais contêm entre 2 e 11 cópias desse domínio.

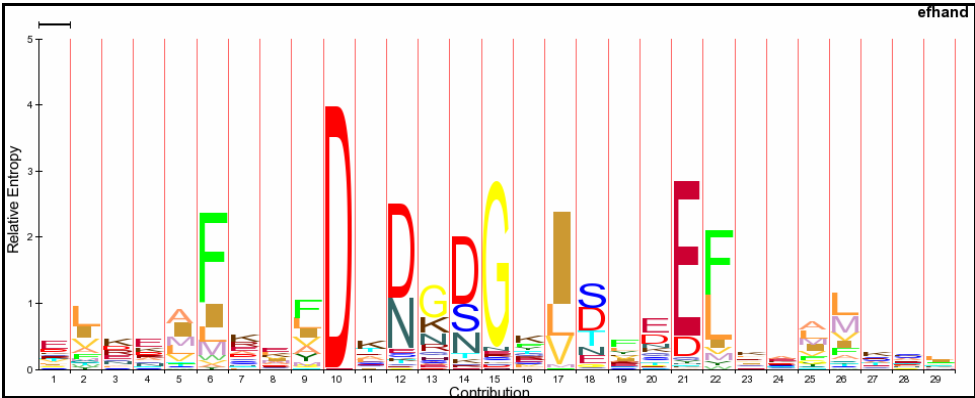


Figura 5.62. Gráfico HMM do domínio EF-hand.
No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

A confiabilidade conjunta da identificação do domínio EF-hand em TcExp02B10 é de $8,8e^{-17}$, sendo que os domínios individuais apresentam uma confiabilidade próxima de $1e10^{-4}$. Eles estão localizados nas posições 60, 97 e 138 da sequência aminoacídica, possuindo 28 aminoácidos cada um (Figura 5.63).

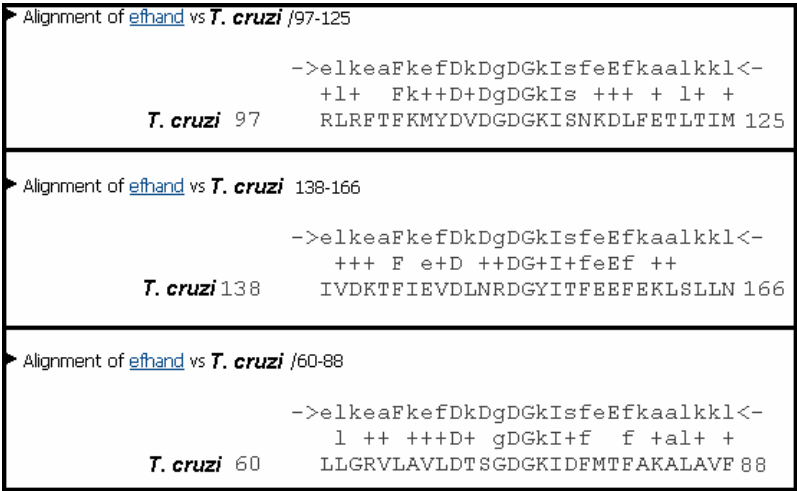


Figura 5.63. Comparação entre a sequência de *T. cruzi* e o domínio EF-hand
Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar a presença dos três domínios já citados. A numeração existente se refere a posição aminoacídica que inicia e termina o domínio dentro do gene em questão.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências com um valor de significância próximo a $1e^{-49}$ para tripanossomatídeos e $1e^{-30}$ para outros organismos selecionados, sendo que esse domínio é conservado de leveduras a humanos. A comparação das proteínas homólogas pode ser vista na Figura 5.64.

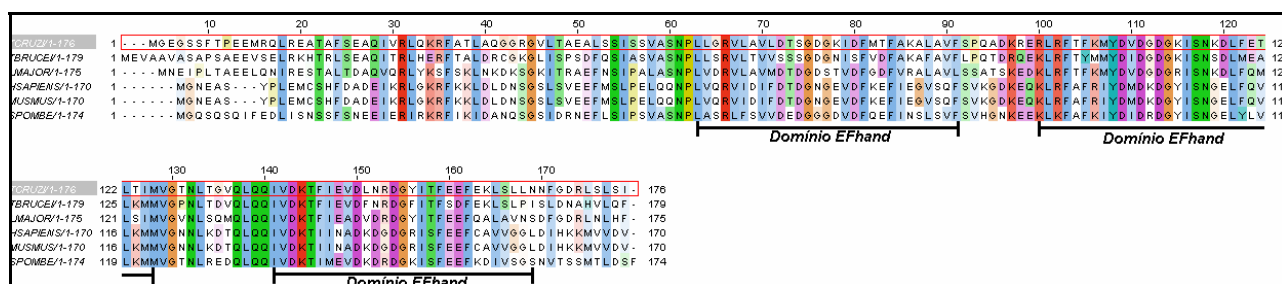


Figura 5.64. Análise de similaridade entre TcExp02B10 e ortólogos de outros eucariotos.
O domínio EFhand está delimitado. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *Western blot* da metaciclogênese pode ser visto na Figura 5.65, onde pode ser observada um aumento de 1,7 vez em aderido 24 horas e uma diminuição de 4,1 vezes em metacíclicos, em relação a epimastigotas. A magnitude desse último resultado deve ser considerada com cuidado, pois não houve sinal em metacíclicos. No entanto, o padrão de diminuição da expressão, de epimastigotas para metacíclicos, é evidente.

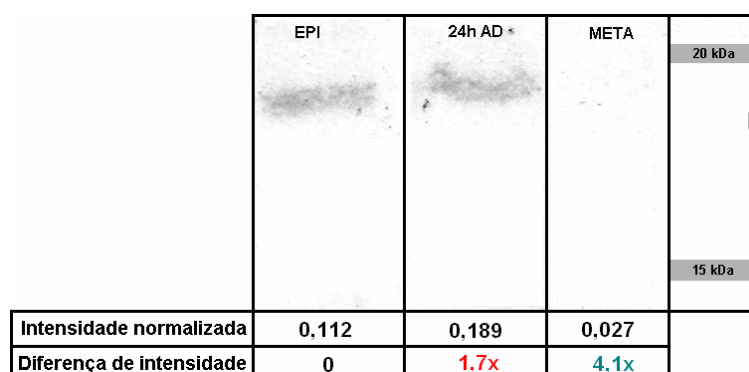


Figura 5.65. Análise de expressão da proteína TcExp02B10 por western blot em extratos de *T. cruzi*.
EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho e verde representam, respectivamente, aumento e diminuição da expressão em relação a epimastigotas

O resultado da imunolocalização de TcExp02B10 pode ser visto na Figura 5.66, onde percebe-se uma localização de fluorescência com menor intensidade na região ao redor do núcleo e grânulos com fluorescência mais intensa no flagelo, na região mais extrema do mesmo.

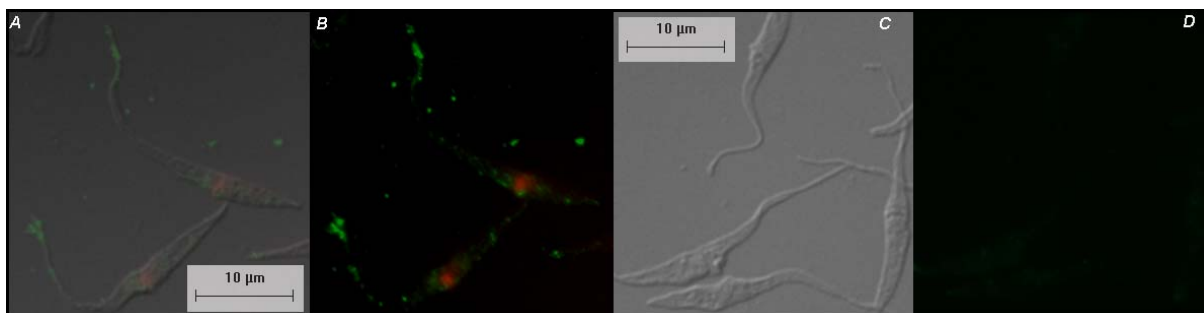


Figura 5.66. Imunolocalização de TcExp02B10.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e a marcação do DNA (DAPI, vermelho); C) contraste diferencial do soro pré-imune. D) Imunofluorescência do soro pré-imune.

5.10.21 TcExp02A06, 8415.t00003, Proteína hipotética conservada

Este gene foi classificado como principal do grupo 6 e possui em *T. cruzi* 2257 pares de bases codificando uma proteína hipotética conservada de 86 kDa, apresentando um domínio denominado PPR. Esse pequeno domínio de 35 aminoácidos está ligado geralmente a RNA, podendo participar de diversas vias, como por exemplo, metabolismo, estabilização e processamento do RNA. Este domínio aparece em grandes quantidades nas proteínas de *A. thaliana* e outros vegetais, porém também é encontrado em outros eucariotos. Possui presença de alguns aminoácidos conservados no domínio, conforme pode ser visto na Figura 5.67.

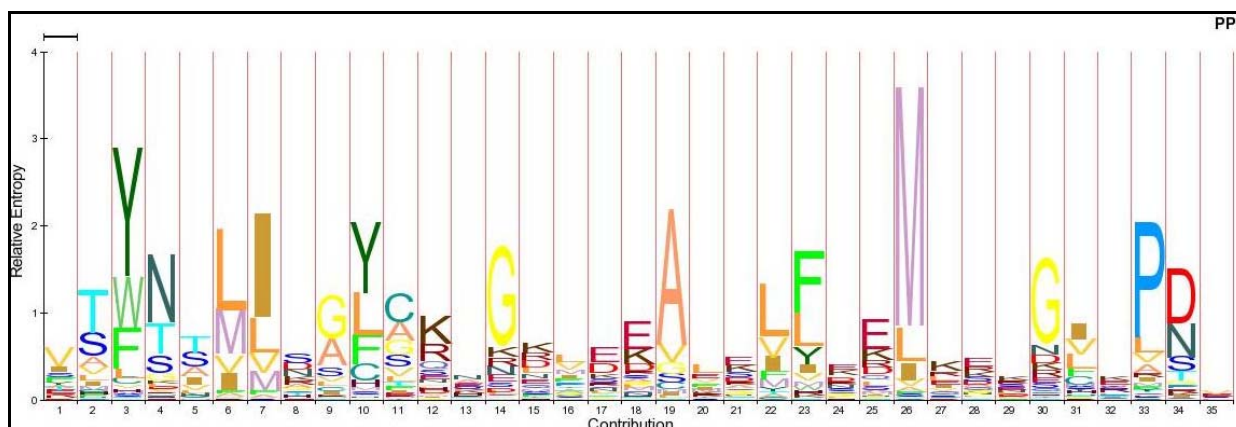


Figura 5.67. Gráfico HMM retirado do programa PFAM mostrando o grau de conservação em cada posição aminoacídica do domínio PPR.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

Quando comparada a similaridade entre o domínio PPR e TcExp02A06 verificou-se um valor de confiabilidade de 0,0011. Os domínios iniciam em TcExp02A06 na posição aminoacídica 69 terminando na posição 103 como pode ser visto na Figura 5.68, evidenciando-se a ocorrência da maioria dos resíduos mais conservados.

Alignment of PPR vs *T. cruzi*

```

PPR ->vtYntlllsgycknGkleeAlelfeeMkekGikPdv<-
      +Y++lI  +++++  A++lf++M+e  G++P
UNKNOWN-T. cruzi 69      LMYTSLIATMGRAKLEWQAYKLF SRMLEQGVRLP 103

```

Figura 5.68. Comparação entre a sequência de *T. cruzi* e o domínio PPR.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 69 da sequência do *T. cruzi* e o término na posição 103.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou sequências com um valor de significância próximo a $1e^{-149}$ somente para tripanossomatídeos, cujo alinhamento pode ser visto na Figura 5.69. É possível evidenciar a alta conservação dessas sequências, mesmo nas regiões além do domínio PPR; outro aspecto que chama a atenção é a grande quantidade de resíduos básicos (K e R) e ácidos (D e E), bem como prolina, que usualmente ocorrem em blocos.

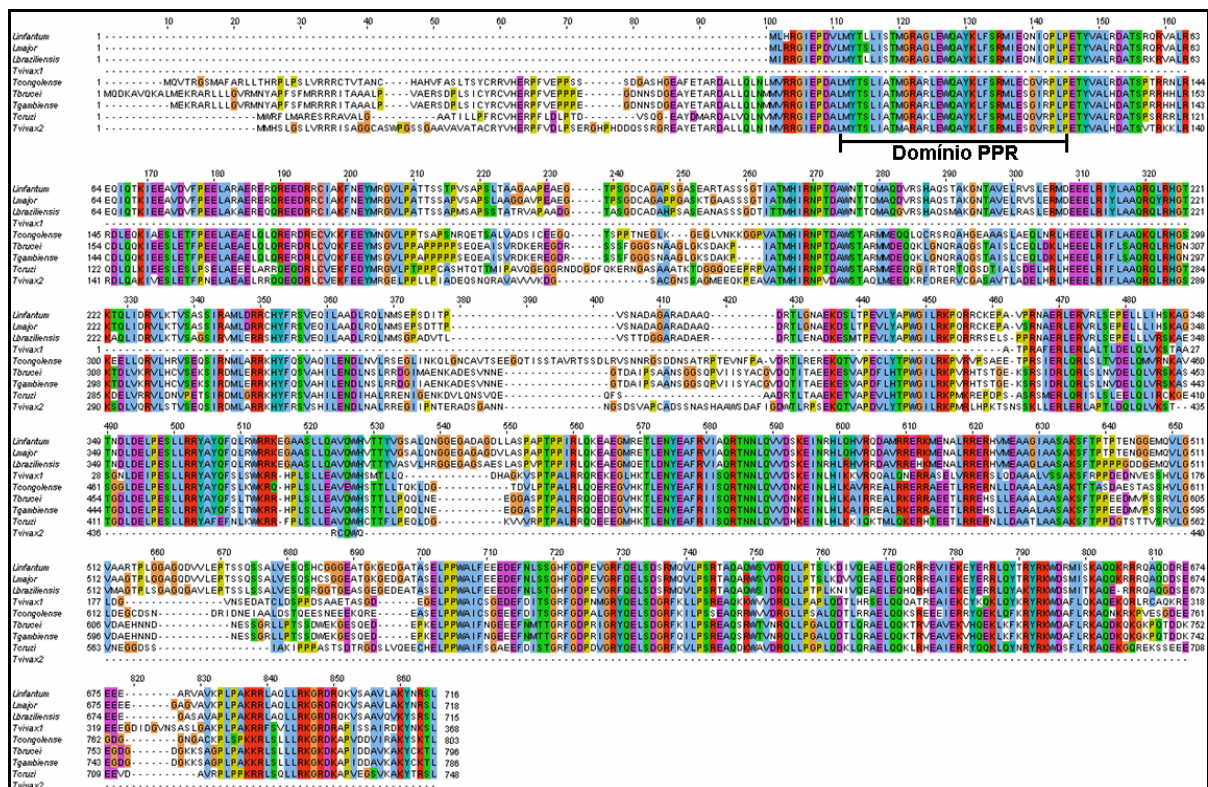


Figura 5.69. Análise de similaridade entre sequências de *T. cruzi* e outros eucariotos.

O Domínio PPR está sendo demarcado abaixo das sequências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém não foi obtida a proteína expressa.

5.10.22 TcExp02B11, 7000.t00002, Proteína hipotética

Este gene está classificado como acessório no gripo 6 e possui em *T. cruzi* 1905 pares de bases codificando uma proteína hipotética de 72,2 kDa. Não possui homólogos em outros tripanossomatídeos e não apresenta domínios pelos bancos de dados PFAM, Interpro e Prosite. Apresenta uma região transmembrana ou peptídeo sinal em sua região amino-terminal e não apresenta ortólogos em outros organismos. No entanto, há um conjunto de proteínas distintas em tripanossomatídeos que são similares a essa sequência, apresentando também uma região transmembrana em sua extremidade aminoterminal e a inexistência de domínios funcionais identificados. O alinhamento protéico dessas sequências pode ser visto Figura 5.70, demonstrando claramente a distinção de TcExp02B11 das demais, o que pode ser melhor evidenciado na árvore filogenética representada na Figura 5.71.

Nessa figura pode se observar a existência de quatro grupos relativos a essa família protéica; sendo que um dos grupos está presente em todos os tripanossomatídeos, incluindo o único representante presente em *Leishmania*; um segundo grupo contém sequências somente de *Trypanosoma*, incluindo dois parálogos em *T. vivax* e os dois últimos grupos são compostos por somente uma proteína de *T. cruzi*, sendo que um deles é representado pela TcExp02B11.

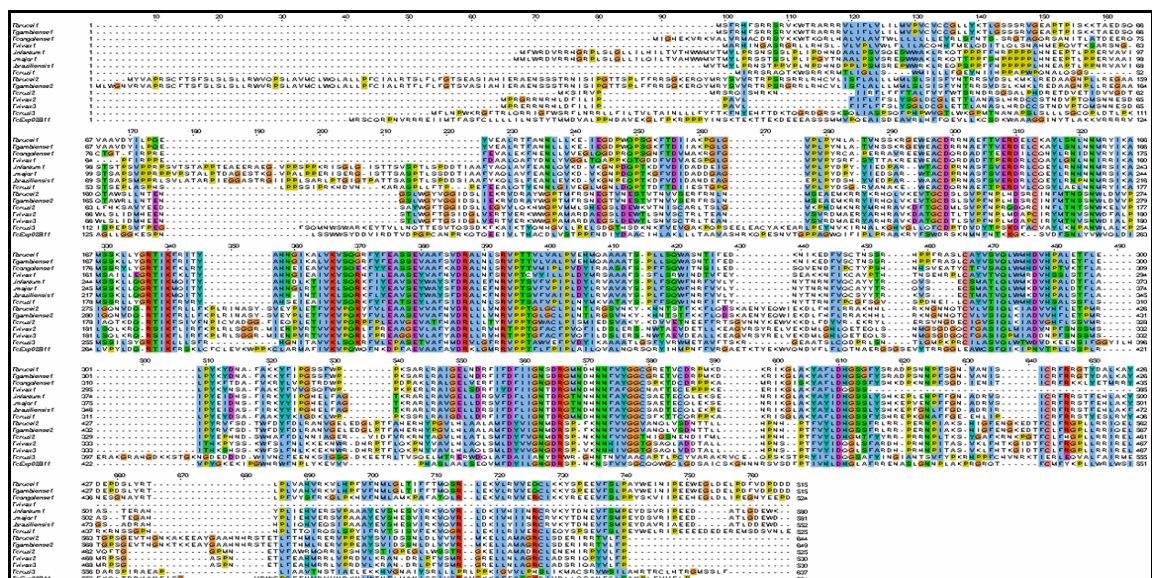


Figura 5.70. Análise de similaridade entre sequências de *T. cruzi* e outros tripanossomatídeos.
Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoácídica

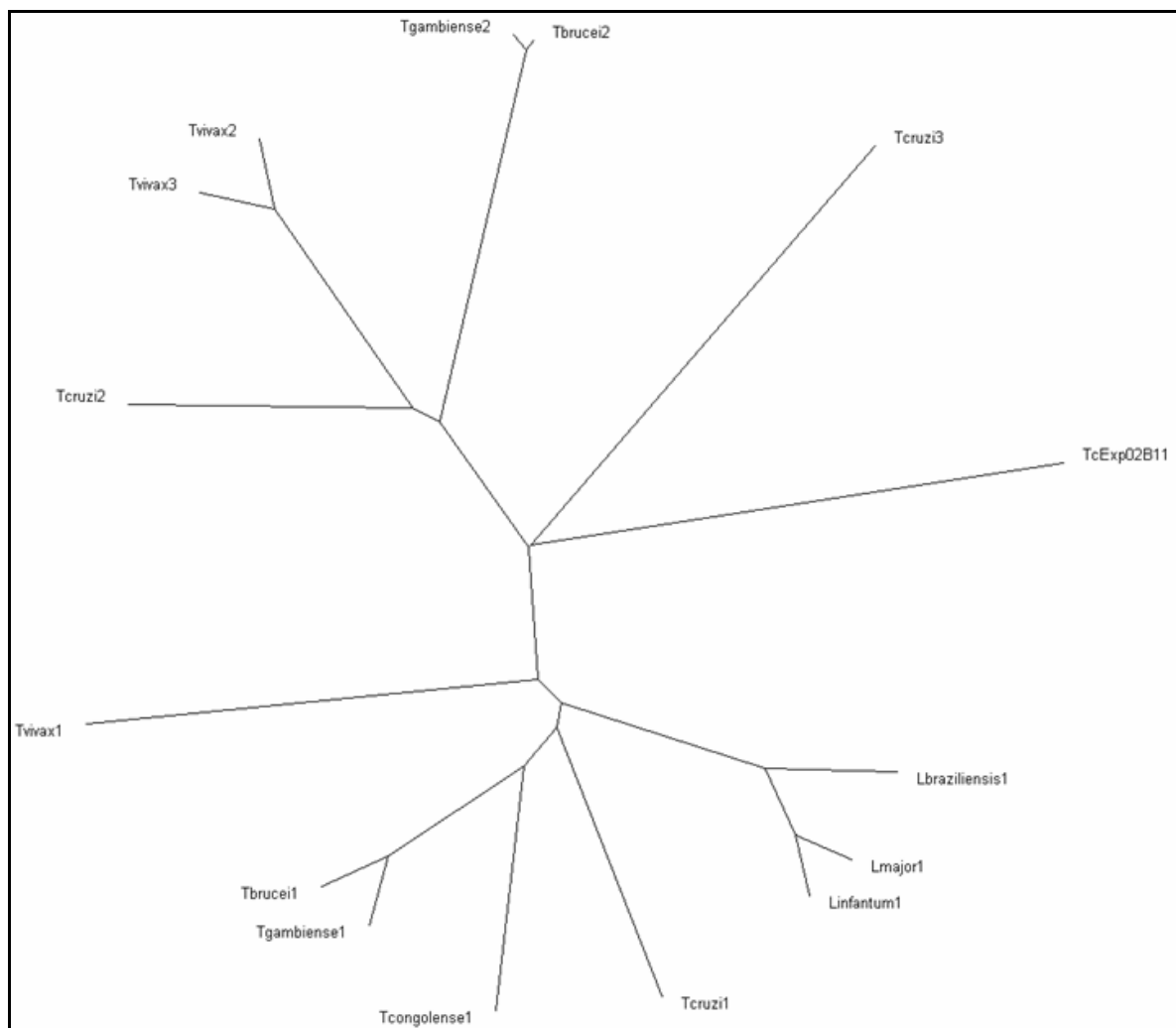


Figura 5.71. Árvore filogenética mostrando as relações evolutivas da família de TcExp02B11.
Quanto maior a distância entre as diferentes espécies menor a conservação das seqüências analisadas.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão, porém não foi obtido proteína expressa em *E. coli*.

5.10.23 TcExp02B12, 8318.t00003, Proteína Hipotética conservada GTPase

Este gene está classificado como secundário no grupo 6 e possui em *T. cruzi* 813 pares de bases codificando uma proteína hipotética conservada de 31,2 kDa, apresentando um domínio denominado ArfGap (ADP-Ribosilation Factor/GTP-ase Activating Protein). Esse domínio PFAM é composto por 137 aminoácidos e possui função de hidrolisar GTP ligado à outra molécula denominada de ARF, a qual está envolvida na regulação do tráfego de membranas e o remodelamento de actina. Além disto, proteínas contendo o domínio em questão parecem estar ligadas a outros processos independentes de ARF, como na formação de vesículas revestidas. Possui como característica a presença de um domínio *zinc-finger* semelhante ao do

tipo GATA, além de outros aminoácidos conservados como asparagina, triptofano, tirosina, histidina, arginina e serina como pode ser verificado na Figura 5.72 (RANDAZZO e HIRSCH; 2003).

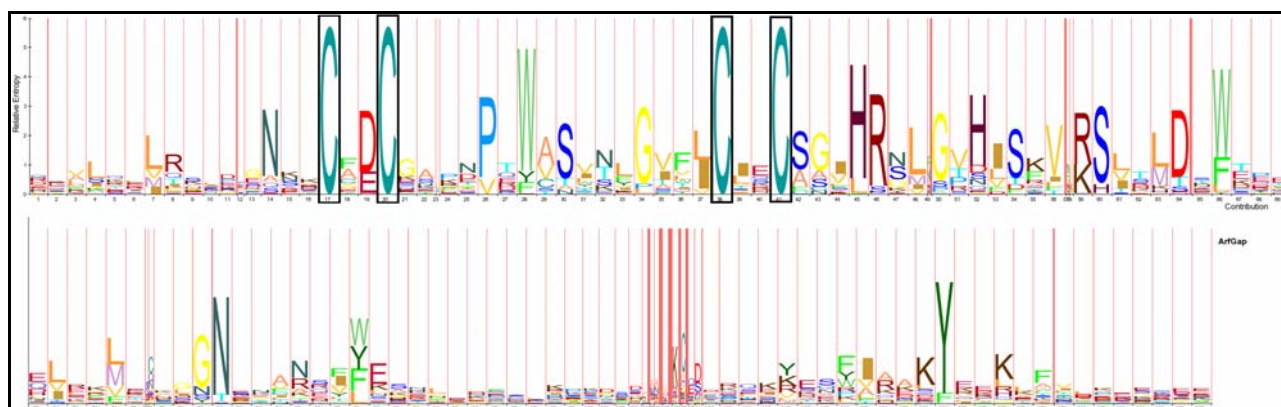


Figura 5.72. Gráfico HMM do domínio ArfGap.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

Quando comparada a similaridade entre o domínio ArfGap e a sequência do *T. cruzi*, verificou-se um valor de confiabilidade aproximadamente a $3,3e^{-40}$. O domínio inicia em TcExp02B12 na posição aminoacídica 47 terminando na posição 168 como pode ser verificado na Figura 5.73.

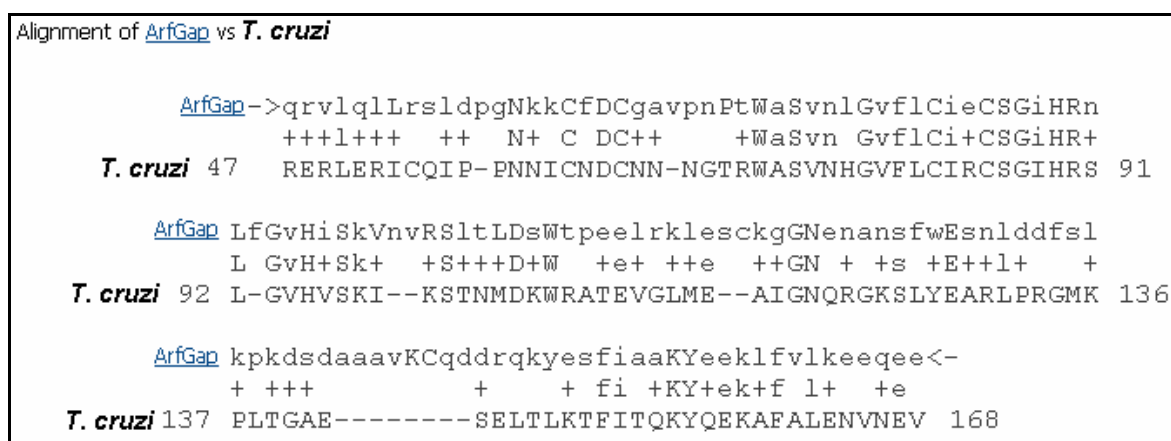


Figura 5.73. Comparação entre a sequência de *T. cruzi* e o domínio ArfGap.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 47 da sequência do *T. cruzi* e o término na posição 168.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências, com um valor de significância próximo a $3e^{-72}$ para *T. brucei*, $2e^{-46}$ para *L. major* e $5e^{-21}$ para diversos outros eucariotos, geralmente na

região contendo somente o domínio. A análise comparativa das seqüências homólogas pode ser vista na Figura 5.74, evidenciando-se a grande conservação do domínio, mesmo em espécies distantes evolutivamente.

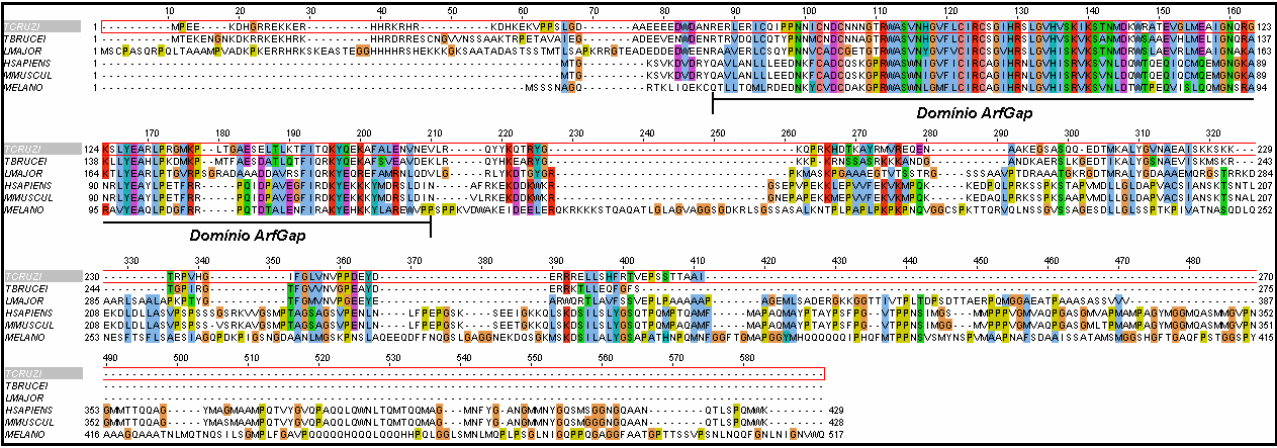


Figura 5.74. Análise de similaridade entre TcExp02B12 e ortólogos de outros eucariotos.
O Domínio ArfGap está sendo demarcado abaixo das seqüências analisadas. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.75, na qual observa-se um forte aumento em aderido 24 horas (35 vezes) e metacíclicos (25 vezes) em relação a epimastigotas. Esses valores de mudança devem ser considerados com cuidado, pois a intensidade do sinal em epimastigotas é muito fraca. De qualquer forma, o padrão de aumento durante a metaciclogênese é evidente.

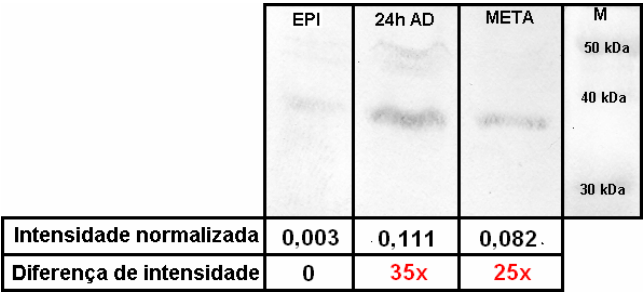


Figura 5.75. Análise de expressão da proteína TcExp02B12 por *western blot* em extratos de *T. cruzi*.
EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho representam aumento da expressão em relação a epimastigotas

Quanto à localização da proteína TcExp02B12 pode-se verificar a sua localização entre o flagelo e o cinetoplasto, conforme visto na Figura 5.76.

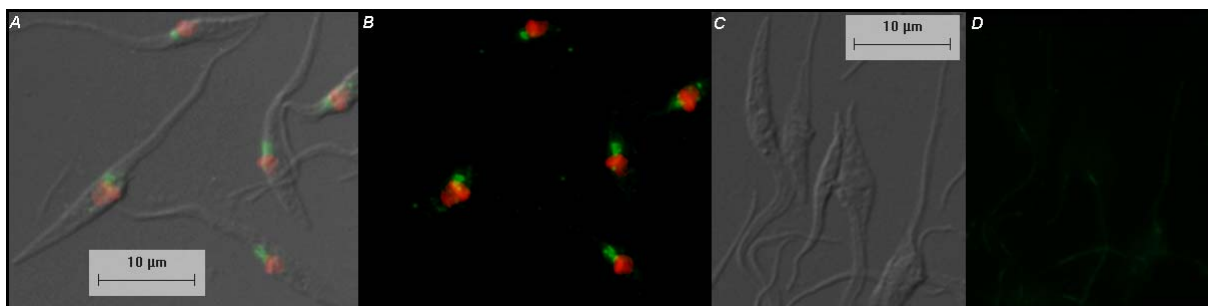


Figura 5.76. Imunolocalização de TcExp02B12

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal específico (verde) e marcação de DNA (DAPI, vermelho); C) contraste diferencial do soro pré-imune. D) Imunofluorescência do soro pré-imune.

5.10.24 TcExp02C01, 7168.t00007, Syntaxin putative

Este gene está classificado como guia do grupo 6 e possui em *T. cruzi* 909 pares de bases codificando uma proteína de 35,5 kDa. Apresenta função anotada como syntaxina possuindo um domínio PFAM denominado SNARE, predito com um valor de confiança igual a $1e^{-20}$ e se estendo por aproximadamente 60 aminoácidos. O domínio está presente normalmente na região carboxi-terminal, possuindo uma região hidrofóbica com a função de ligação em bicamadas lipídicas, podendo estar presente em diferentes membranas e compartimentos, indicando sua especificidade para os eventos de fusão de membranas.

As proteínas da família SNARE são divididas basicamente em dois tipos, v-SNAREs e t-SNAREs. Ambas participam do sistema de fusão de membranas, sendo que as proteínas presentes nas vesículas são denominadas de v-SNAREs e as presentes na organela ou membrana plasmática de t-SNAREs (de alvo, *target*). A syntaxina é enquadrada na classe de proteínas t-SNAREs. Existe uma segunda nomenclatura proposta baseada em um único resíduo presente no domínio que pode ser uma glutamina (Q-SNARE) ou uma arginina (R-SNARE). Na maioria dos casos as proteínas presentes nas vesículas são do tipo R e na membrana alvo do tipo Q (WEIMBS *et al*, 1997; FASSHAUER *et al*, 1998) Em TcExp02C01, o domínio PFAM predito apresenta uma glutamina na posição 30 do domínio, caracterizando-o como uma proteína do tipo Q-SNARE, mas também possui uma conservação acentuada de leucina e isoleucina como pode ser visto na Figura 5.77.

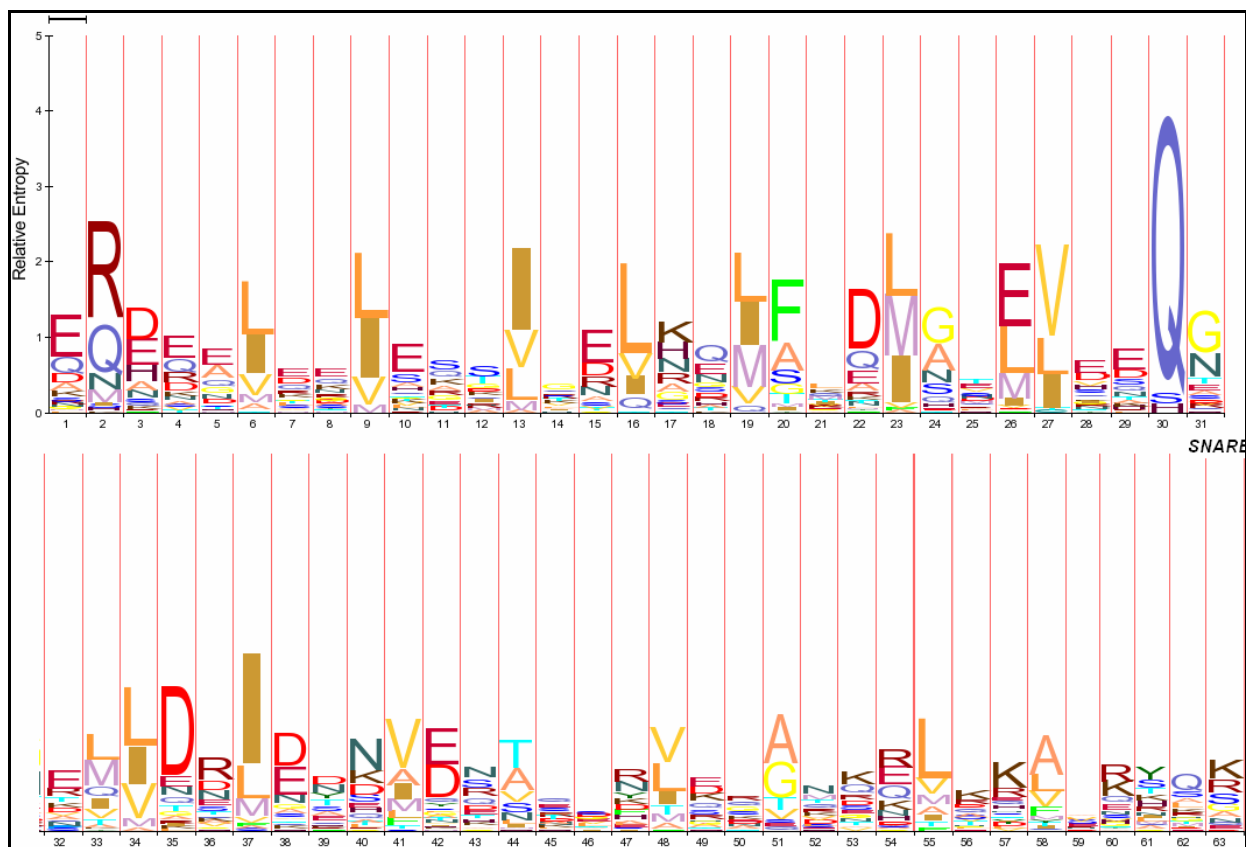


Figura 5.77. Gráfico HMM do domínio SNARE.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.

Quando comparada a proteína TcExp02C01 com o domínio SNARE, verifica-se que a região de similaridade se inicia na posição aminoacídica 212 e termina na posição 274, podendo-se notar a preservação do aminoácido glutamina na posição 30 do domínio como pode ser verificado na Figura 5.78.

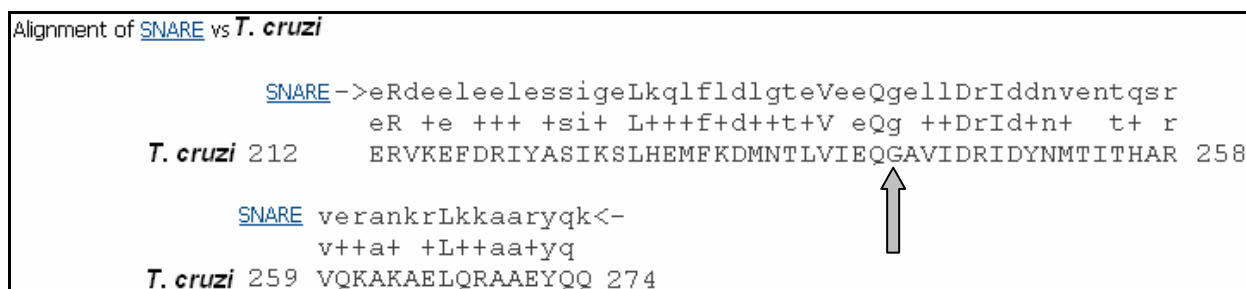


Figura 5.78. Comparação entre a sequência de *T. cruzi* e o domínio SNARE.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 212 da sequência do *T. cruzi* e o término na posição 274. Seta indicando a glutamina da posição 30.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências similares com um valor de significância

próximo a $1e^{-100}$ para tripanossomatídeos e aproximadamente $1e^{-17}$ para outros organismos eucarióticos, sendo diversas anotadas como sintaxina 16. Embora a similaridade com as seqüências de organismos eucarióticos mais distantes seja relativamente baixa, a região similar se estende por praticamente toda a extensão da proteína, mesmo em regiões fora do domínio PFAM, conforme visto na Figura 5.79.

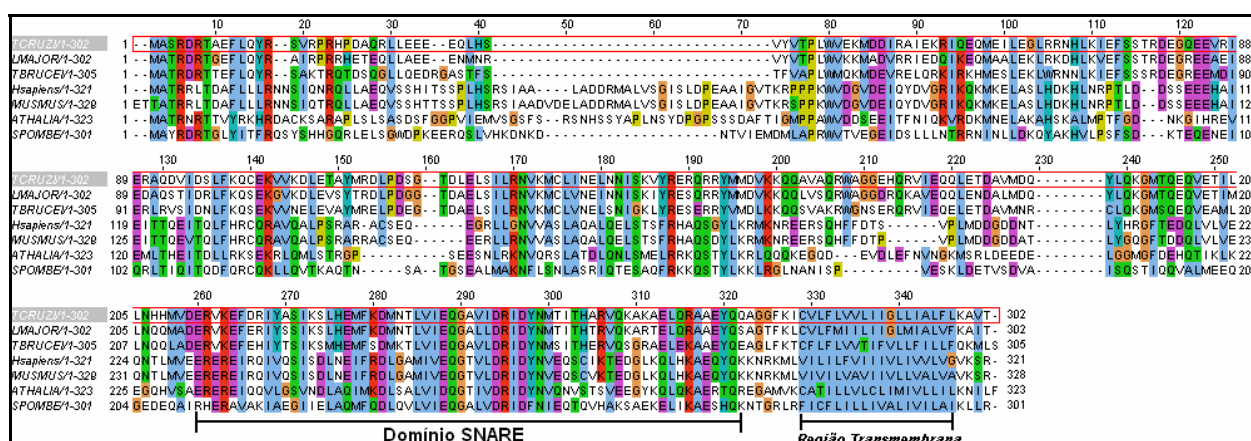


Figura 5.79. Análise de similaridade entre seqüências de *T. cruzi* e outros eucariotos. O Domínio Snare está sendo demarcado abaixo das seqüências analisadas, bem como a região transmembrana característica. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.80, observando-se um aumento claro de 3,2 vezes em metacíclicos, em comparação com epimastigotas.

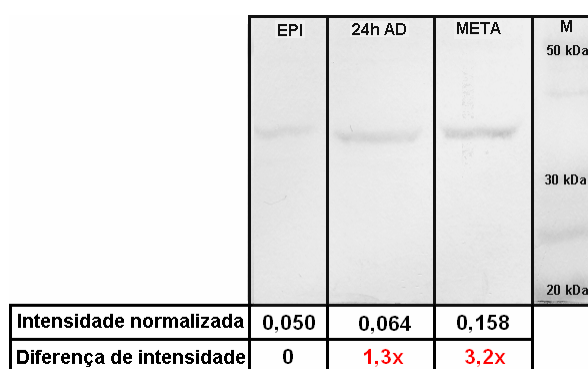


Figura 5.80. Análise de expressão da proteína TcExp02C01 por *western blot* em extratos de *T. cruzi*. EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho representam aumento da expressão em relação a epimastigotas

A imunofluorescência da proteína TcExp02C01 a localiza bem definidamente em uma região próxima ao cinetoplasto estando entre este e o flagelo, como pode ser verificado na Figura 5.81.

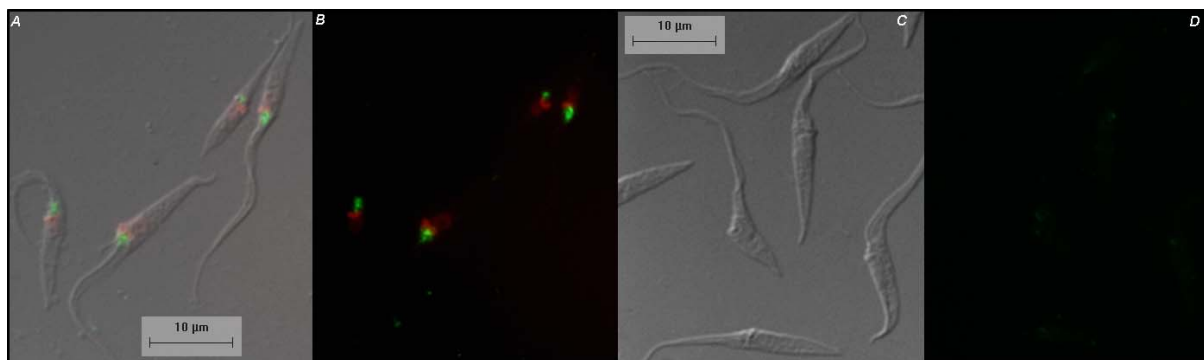


Figura 5.81. Imunolocalização da proteína relativa ao gene 7168.t00007.

A) contraste diferencial de fase com sobreposição da figura B. B) Imunolocalização utilizando o anticorpo policlonal anti-7168.t00007. Em verde a marcação referente ao anticorpo específico e em vermelho a marcação de DNA utilizando DAP. C) contraste diferencial de fase relativo ao soro pré-imune. D) Imunofluorescência obtida com a utilização do soro pré-imune.

5.10.25 TcExp02A07, 5323.t00004, Proteína hipotética conservada

Este gene está classificado como principal do grupo 7 e possui em *T. cruzi* 1362 pares de bases codificando uma proteína de 51 kDa, a qual apresenta um domínio denominado PKinase. Conforme mencionado anteriormente, esse domínio faz parte de um extenso grupo de proteínas que compartilham um domínio catalítico conservado em serina/treonina e tirosinas quinases tendo como características a presença de algumas glicinas conservadas na região amino-terminal do domínio próximo a um resíduo de lisina, também conservado, o qual parece estar associado à ligação de ATP. Além destes resíduos existem resíduos de ácido aspártico conservados na região central do domínio sendo importantes para a atividade catalítica da enzima, os quais podem ser vistos na Figura 5.24.

Quando comparado o domínio PKinase e a sequência TcExp02A07, verificou-se um valor de confiabilidade de $6e^{-90}$. O início do domínio na sequência gênica relativa ao *T. cruzi* ocorre na posição 13 e termina na posição 302 como mostra a Figura 5.82.



Figura 5.82. Comparação entre TcExp02A07 e o domínio PKinase.

Entre as duas seqüências encontra-se a conservação aminoacídica. Pode-se verificar o início do domínio na posição 13 da seqüência do *T. cruzi* e o término na posição 302.

A comparação da seqüência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias seqüências, com um valor de significância em geral muito alto, mesmo para organismos mais distantemente relacionados, devido ao fato de que o domínio é extenso e relativamente conservado, conforme pode ser visto na Figura 5.83. De maneira geral, a similaridade se restringe ao domínio e as proteínas mais similares estão geralmente anotadas como uma quinase ativada por mitógeno (MAPK).

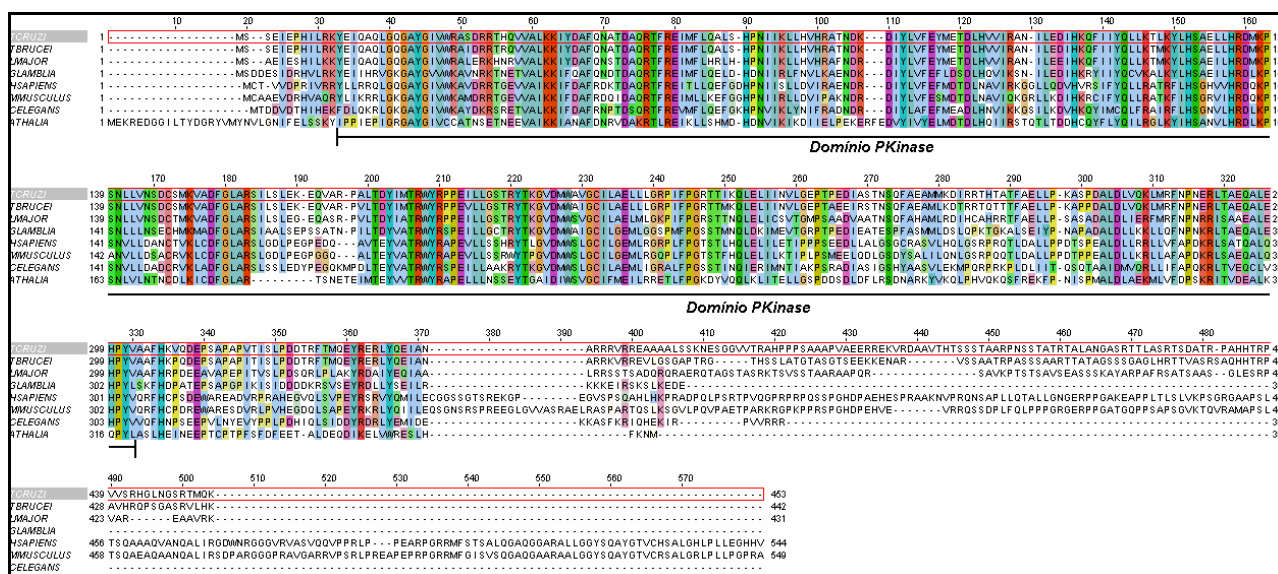


Figura 5.83. Análise de similaridade TcExp02A07 e ortólogos de outros eucariotos.

O domínio PKinase está delimitado. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado obtendo-se o vetor de entrada e de expressão. A expressão foi obtida em baixa quantidade porém não foi possível obtê-la purificada, devido à perda ao longo do processo de purificação.

5.10.26 TcExp02C02, 8359.t00034, Proteína Hipotética Conservada

Este gene está classificado como acessório no grupo 7 e possui em *T. cruzi* 1086 pares de bases codificando uma proteína hipotética conservada de 39,6 kDa, a qual apresenta um domínio denominado TLD. Esse domínio não possui função conhecida, tendo sido proposto por DOERKS *et al*, 2003 e postula-se que possua função enzimática. O domínio PFAM predito possui 165 aminoácidos e apresenta alguns aminoácidos muito conservados como pode ser visto na Figura 5.84.

Quando comparada a similaridade entre o domínio TLD e a sequência do *T. cruzi* TcExp02A07, o programa PFAM fornece dois alinhamentos, um com valor de significância $9e-07$ contendo 54 aminoácidos e um maior com 205 aminoácidos com um valor de $4.5e-06$ englobando o primeiro já citado. O início do domínio na sequência gênica relativa ao *T. cruzi* ocorre na posição 75 e termina na posição 307 como mostra a Figura 5.85

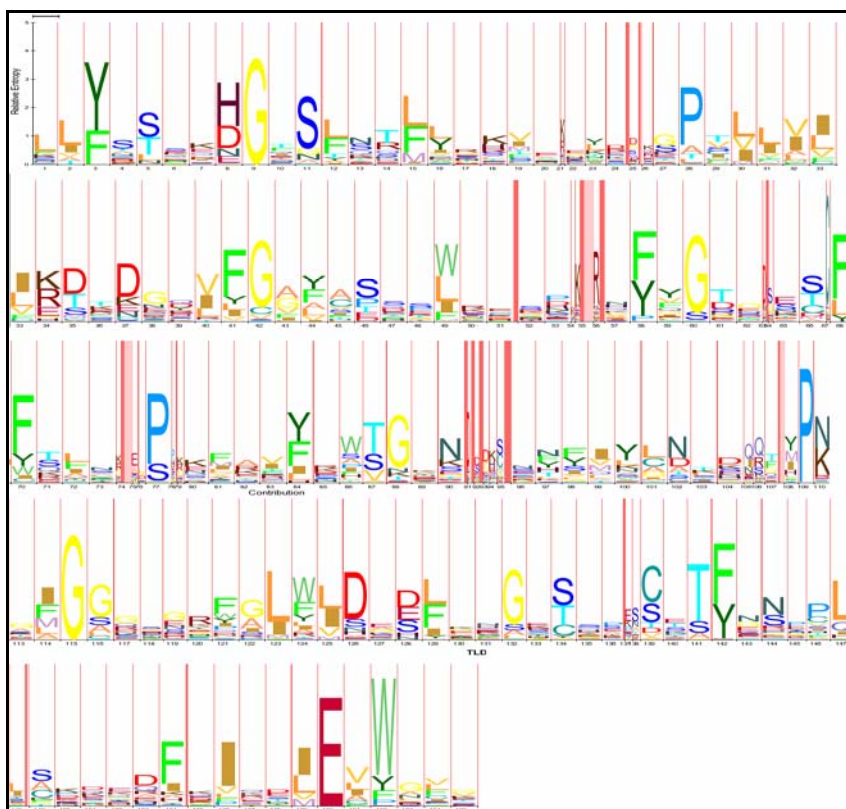


Figura 5.84. Gráfico HMM do domínio TLD.

No eixo inferior do gráfico estão numeradas as posições aminoacídicas do domínio e no eixo lateral a entropia relativa.



Figura 5.85. Comparação entre a TcExp02C02 e o domínio TLD.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp identificou poucas sequências similares, somente em tripanossomatídeos. A comparação das sequências homólogas pode ser vista na Figura 5.86.

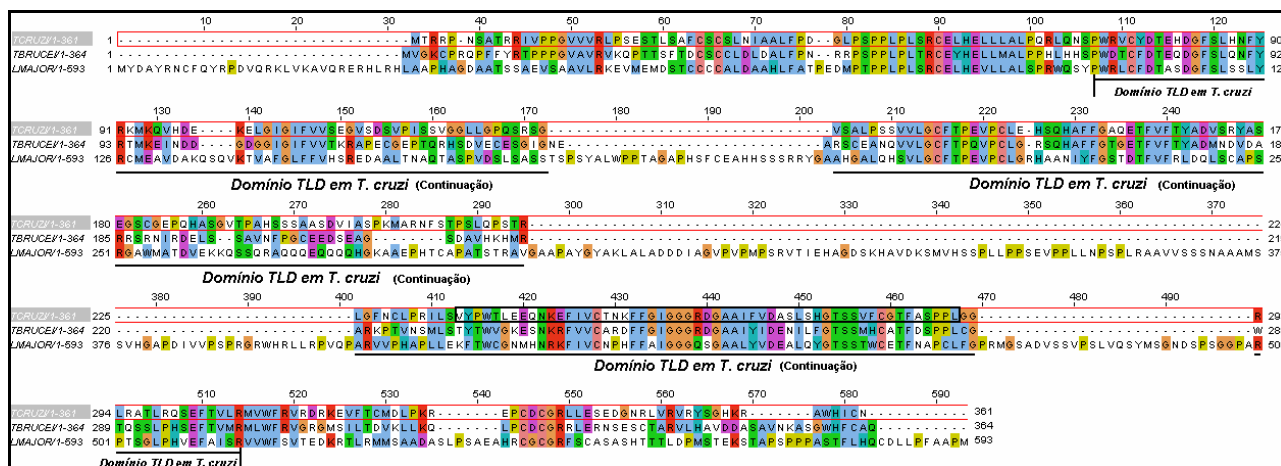


Figura 5.86. Análise de similaridade entre TcExp02C02 e ortólogos de outros tripanossomatídeos.

O domínio TLD está delimitado. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado, obtendo-se o vetor de entrada e de expressão. A expressão da proteína foi obtida, bem como o soro policlonal, porém este só reconheceu a proteína recombinante, não reagindo com o extrato de *T. cruzi*.

5.10.27 TcExp02C03, 7741.t00008, Proteína Hipotética Conservada

Este gene está classificado como secundário no grupo 7 e possui em *T. cruzi* 1209 pares de bases codificando uma proteína hipotética conservada de 47,7 kDa, e não apresenta domínios preditos na comparação com os bancos de dados PFAM ou Interpro. Só foi encontrado um gene similar em *T. brucei* com um valor de significância relativamente alto, $2e^{-48}$. Não foi identificado um gene similar em outros Trypanosomas e em Leishmania. A análise das seqüências homólogas pode ser vista na Figura 5.87

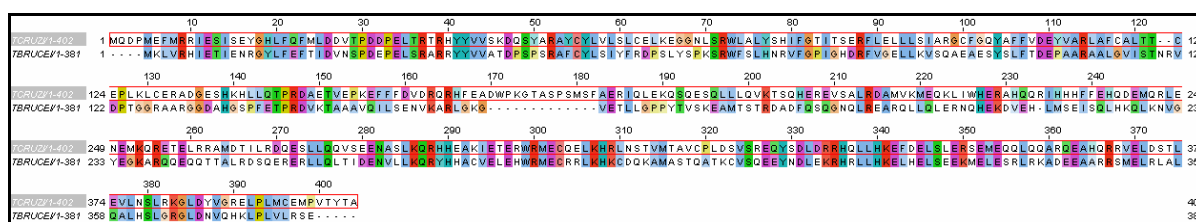


Figura 5.87. Análise de similaridade entre seqüências de *T. cruzi* e *T. cruzi*.

Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Este gene foi clonado, obtendo-se o vetor de entrada e de expressão. A expressão da proteína foi obtida, bem como o soro policlonal, porém este só reconheceu a proteína recombinante, não reagindo com o extrato de *T. cruzi*.

5.10.28 TcExp02C04, 6996.t00038, Centrin Putative

Este gene está classificado como guia do grupo 7 e possui em *T. cruzi* 546 pares de base codificando uma proteína de 21 kDa, apresentando três domínios denominados EF-hand. Esse pequeno domínio de 29 aminoácidos faz parte de um extenso grupo de proteínas que possuem como função a ligação ao cálcio. A representação desse domínio, com suas posições mais conservadas pode ser vista na Figura 5.62.

As centrininas são proteínas pertencentes à grande família das proteínas ligadoras de cálcio e fazem parte do centrosomo, que estão associados aos centríolos, sendo uma das poucas proteínas centríolares identificadas. São encontradas em todos eucariotos e são necessárias para a duplicação dos centríolos, podendo estar envolvidas em outros processos celulares, como excisão flagelar, exportação de mRNA nuclear e recombinação homóloga (ZAMORA & MARSHALL, 2005).

A confiabilidade combinada da identificação dos domínios EF-hand é de $5e^{-24}$, demonstrando confiabilidade adequada nas semelhanças das seqüências. Os domínios têm seus inícios em *T. cruzi* nas posições 38, 114 e 150 da seqüência aminoacídica de

TcExp02C04, possuindo cada um deles 28 aminoácidos cada um dos 3 domínios, como pode ser observado na Figura 5.88.

► Alignment of <i>efhand</i> vs <i>T. cruzi</i>	
<i>efhand</i>	->elkeaFkefDkDgDGkIsfeEfkaalkkl<- e+teaF++fD D++G I++ E+k +++l
<i>T. cruzi</i>	38 EIREAFELFDSDKNGLIDVHEMKVSMRAL 66
► Alignment of <i>efhand</i> vs <i>T. cruzi</i>	
<i>efhand</i>	->elkeaFkefDkDgDGkIsfeEfkaalkkl<- e+ +aF++fD ++ GkIs+ ++++ ++l
<i>T. cruzi</i>	114 EMIKAFQLFDENNTGKISLRSLLRRVAREL 142
► Alignment of <i>efhand</i> vs <i>T. cruzi</i>	
<i>efhand</i>	->elkeaFkefDkDgDGkIsfeEfkaalkkl<- el+++++eFD+D+DG+I+++Ef+a++ +
<i>T. cruzi</i>	150 ELQAMIDEFDTDQDGEINLDEFLAIMLED 178

Figura 5.88. Comparação entre a sequência de *T. cruzi* e o domínio EF-hand.

Entre as duas sequências encontra-se a conservação aminoacídica. Pode-se verificar a presença dos três domínios já citados. A numeração existente se refere a posição aminoacídica que inicia e termina o domínio dentro do gene em questão.

A comparação da sequência de *T. cruzi* com as proteínas do Genbank, utilizando o programa BLASTp, identificou várias sequências, com um valor de significância igual a $3e^{-73}$ para *T. brucei*, $1e^{-60}$ para *L. major*, e aproximadamente $1e^{-36}$ para outros organismos selecionados. A análise das sequências homólogas pode ser vista na Figura 5.89, evidenciando-se a grande conservação, mesmo em regiões externas aos domínios.

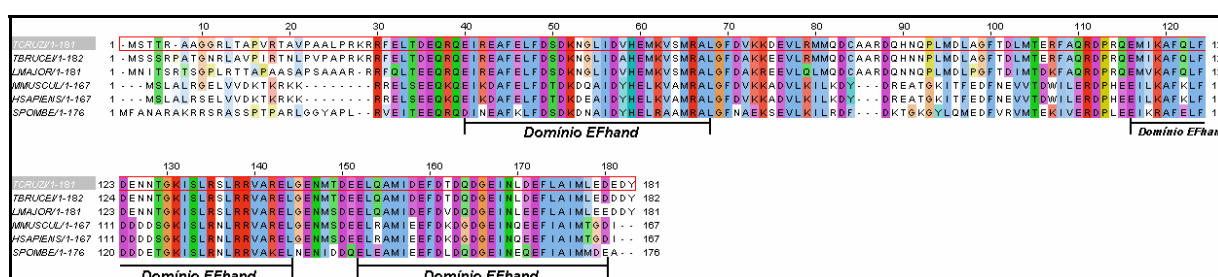


Figura 5.89. Análise de similaridade entre TcExp02C04 e ortólogos de outros eucariotos.

O domínio EFhand está delimitado. Na parte superior dos alinhamentos encontram-se as numerações relativas à posição aminoacídica.

Para este gene todas as etapas laboratoriais foram realizadas com sucesso. O resultado da análise por *western blot* da metaciclogênese pode ser visto na Figura 5.90, observando-se um aumento pouco pronunciado em aderido 24 horas e uma diminuição também moderada em metacíclicos..

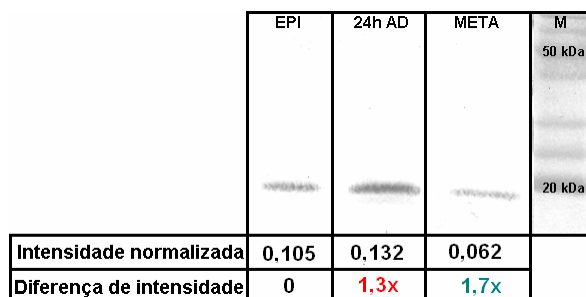


Figura 5.90. Análise de expressão da proteína TcExp02C04 por western blot em extratos de *T. cruzi*.
EPI: epimastigota; 24 AD: intermediária de diferenciação; META: metacíclico; M: Marcador de peso molecular. Os valores em vermelho e verde representam, respectivamente, aumento e diminuição da expressão em relação a epimastigotas.

A localização da proteína TcExp02C04 aparece bem definida, estando co-localizada com o cinetoplasto, como pode ser visto na Figura 5.91.

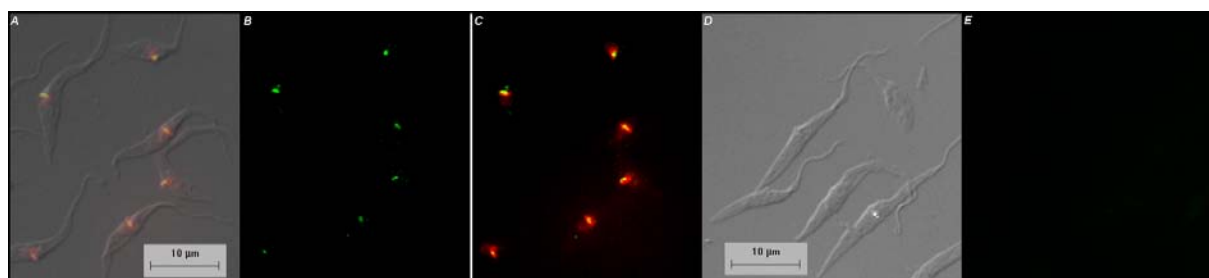


Figura 5.91. Imunolocalização de TcExp02C04.

A) contraste diferencial de fase com sobreposição da figura C. B) Imunolocalização utilizando o anticorpo policlonal específico (verde). C) Sobreposição das imagens de fluorescência (vermelho, DAPI); D) contraste diferencial do soro pré-imune. E) Imunofluorescência do soro pré-imune.

5.11 Consolidação dos resultados laboratoriais

O número de proteínas avaliadas foi relativamente grande e a análise é composta por diversas etapas, tornando complexo todo o processo de produção. Conforme mencionado acima, as etapas de expressão protéica e produção de anti-soros representaram os principais gargalos em toda a análise. No entanto, isso representa aspectos gerais de todo o processo e, devido à diversidade de elementos analisados, é possível esmiuçar melhor os padrões observados. Na Tabela 5.4, está representada a eficiência para cada proteína e para cada etapa estudada.

Tabela 5.4. Eficiência de cada etapa de análise para as 28 proteínas analisadas.

ID IBMP	Classificação	Tamanho	Módulos (Domínios)	C	E	P	S	W	I
TcExp02A01	Gene Principal	693 nt, 25 kDa	ALBA						
TcExp02A08	Gene Acessório	1368 nt, 52 kDa	Hypothetical protein						
TcExp02A09	Gene Secundário	1197 nt, 43 kDa	Saccharopine dehydrogenase						
TcExp02A10	Gene Guia	1080 nt, 41 kDa	Sterol 24-c-methyltransferase, putative						
TcExp02A02	Gene Principal	3126 nt, 117 kDa	Protein kinase						
TcExp02A11	Gene Acessório	1644 nt, 61 kDa	Hypothetical protein						
TcExp02A12	Gene Secundário	1320 nt, 49 kDa	Zinc finger C2H2-type						
TcExp02B01	Gene Guia	681 nt, 24 kDa	Ras-related protein rab-5, putative						
TcExp02A03	Gene Principal	3027 nt, 114 kDa	FKBP-type						
TcExp02B02	Gene Acessório	1089 nt, 40 kDa	Hypothetical protein						
TcExp02B03	Gene Secundário	954 nt, 35 kDa	Zinc finger, C2H2 type						
TcExp02B04	Gene Guia	1095 nt, 41 kDa	DREV methyltransferase, putative						
TcExp02A04	Gene Principal	1353 nt, 51 kDa	SWIM zinc finger						
TcExp02B05	Gene Acessório	1485 nt, 53 kDa	Hypothetical protein						
TcExp02B06	Gene Secundário	249 nt, 9 kDa	CSL zinc finger						
TcExp02B07	Gene Guia	354 nt, 13 kDa	Dynein light chain 2B, putative						
TcExp02A05	Gene Principal	606 nt, 22 kDa	ASF1-like						
TcExp02B08	Gene Acessório	516 nt, 20 kDa	Hypothetical protein						
TcExp02B09	Gene Secundário	1035 nt, 38 kDa	Programmed cell death protein 2						
TcExp02B10	Gene Guia	531 nt, 20 kDa	Calcineurin B subunit, putative						
TcExp02A06	Gene Principal	2247 nt, 86 kDa	Signal peptide/PPR Repeat						
TcExp02B11	Gene Acessório	1905 nt, 72 kDa	Hypothetical protein						
TcExp02B12	Gene Secundário	963 nt, 37 kDa	GTP-ase activating protein, Arf						
TcExp02C01	Gene Guia	909 nt, 36 kDa	Syntaxin, putative						
TcExp02A07	Gene Principal	1362 nt, 51 kDa	Kinase						
TcExp02C02	Gene Acessório	1086 nt, 40 kDa	TLD						
TcExp02C03	Gene Secundário	1209 nt, 48 kDa	Hypothetical protein conserved						
TcExp02C04	Gene Guia	546 nt, 21 kDa	Centrin, putative						

C – clonagem; E- expressão; P- purificação; S – imunização; W – western; I - imunolocalização

Em nenhum dos sete grupos foi concluído o fluxograma para as quatro proteínas que o constituem. Esse resultado havia sido previsto, sendo o grande motivador de se estabelecer as medidas de redundância.

Somente o grupo 1 apresentou a obtenção de resultados completos para três dos quatro genes. O gene acessório (proteína hipotética) foi o único não concluído, não se obtendo sua expressão em *E. coli*. Inicialmente, poderia se supor que, pelo fato de ser classificada como uma proteína hipotética, representaria um erro de predição e a região codificadora fictícia não produziria aminoácidos viáveis. No entanto, foi possível identificar um ortólogo em *L. infantum*, o que enfraqueceu essa hipótese.

Somente o grupo 6 apresentou a obtenção de resultados completos para dois dos quatro genes analisados. Quatro grupos (2, 4, 5 e 7) obtiveram resultados completos para somente um dos quatro genes, sempre o gene guia. O grupo 3 não obteve resultado completo para nenhum de seus genes.

Existe uma diversidade de padrões de sucesso e parâmetros intrínsecos às proteínas. No entanto, chama a atenção a associação entre obtenção de resultados completos e classificação do gene como guia, isto é, aquele cuja função está melhor estabelecida. De um total de sete genes guias, obteve-se resultados totais para seis (85,7%), sendo que a única falha foi devida à morte dos dois camundongos inoculados com a proteína TcExp02B04. De um total de 21 genes codificadores de proteínas hipotéticas, obteve-se resultados totais para somente três (14,3%). Uma análise pelo teste exato de Fisher mostrou que a probabilidade dessa proporção ocorrer por acaso é de somente 0,1%.

Diversos fatores podem justificar essa associação. Um fator diretamente relacionado à probabilidade de se obter uma reação do anticorpo em *western* e imunofluorescência (etapas finais do fluxograma e necessárias para considerar como se obtendo resultados completos) é o nível de expressão protéica proporcional. A princípio, genes com função conhecida são constituintes de vias biológicas mais estudadas e, provavelmente, mais necessárias e expressas em uma célula. Por outro lado, proteínas com função desconhecida (hipotéticas conservadas) podem, a princípio, terem uma expressão menor, dificultando a identificação em *western blot* e imunofluorescência. Esse ponto será avaliado futuramente ao se utilizar quantidades maiores de extrato protéico de *T. cruzi* na análise por *western* e também pela realização da imunofluorescência em tripomastigotas, já que a maioria das proteínas hipotéticas analisadas apresenta um padrão de aumento durante a metaciclogênese.

A etapa de expressão protéica foi o maior gargalo do presente trabalho. No entanto, a hipótese sugerida acima para explicar a discrepância entre os grupos dos genes guia codificadores de proteínas hipotéticas não deve influenciar negativamente as proteínas hipotéticas com relação à sua expressão em sistema heterólogo.

Com o intuito de avaliar essa hipótese, foi verificado que se obteve expressão para sete proteínas dos genes guias (100%) e 11 dos genes codificadores de proteínas hipotéticas (52,4%), sendo o valor de p calculado pelo teste exato de Fisher de 2,7%. Isso indica que já na etapa de expressão há uma perda significativa estatisticamente de proteínas hipotéticas, em relação às proteínas guias, o que indica que, além da hipótese sugerida acima, de nível de expressão endógena da proteína em *T. cruzi*, outros fatores estão dificultando a obtenção de

resultados, nesse caso a expressão, para as proteínas hipotéticas. Uma possibilidade é o fato de que na verdade essa proteína não ocorra naturalmente, representando um erro de predição, e a tentativa de expressá-la em *E. coli* gere um polipeptídeo aberrante, o qual seria ou degradado rapidamente ou letal à célula.

Os dados acima citados definem claramente uma associação entre eficiência do processo de caracterização funcional e o grau de conhecimento sobre a função do gene. Antes de podermos realizar uma análise mais profunda das causas relativas a esse perfil, é necessário repetir futuramente os protocolos de expressão, *western blot* e imunolocalização. Isso se deve ao fato de que para o presente trabalho definiu-se uma estratégia de avanço constante. Portanto, é essencial testar outros métodos, para as etapas supracitadas, a fim de se obter um conjunto mais amplo de dados, os quais serão analisados em detalhes para identificar possíveis fatores determinantes da potencialidade de caracterização funcional.

5.12 Análise comparada do transcriptoma e proteômica de *T. cruzi*

Na Figura 5.92, é possível verificar os padrões de expressão diferencial avaliados pelo microarranjo de DNA (transcriptoma) e pela quantificação baseada em *western blot* e anticorpos policlonais (proteoma). Os valores relativos à expressão diferencial protéica são relativos à epimastigotas e, portanto, a coluna relativa a essa amostra apresenta sempre um sinal de não modulação; os valores relativos à expressão diferencial do mRNA também são relativos a epimastigota, mas consistem em uma classificação discreta do padrão geral de modulação nas amostras avaliadas (epimastigotas em estresse, em diferenciação por 3, 12 e 24 horas e metacíclicos) do experimento de metaciclogênese denominado Met16.

A classificação realizada nos dados de microarranjo favorece a representação da tendência geral de modulação, diminuindo-se a complexidade inerente a esse processo de quantificação. De maneira geral, os padrões de modulação que ocorrem nos níveis de mRNA durante a metaciclogênese são relativamente monotônicos, conforme pode ser visto na Figura 4.4 a Figura 4.10, e a representação de uma tendência geral única é, geralmente, compatível com todos os pontos avaliados.

Das 23 proteínas quantificadas pelo uso de anti-soro policonal, somente uma não apresenta sonda no microarranjo de DNA, a GTPase TcYchF, a qual mostrou um padrão de diminuição forte em sua expressão protéica na fase metacíclica, em comparação com epimastigotas.

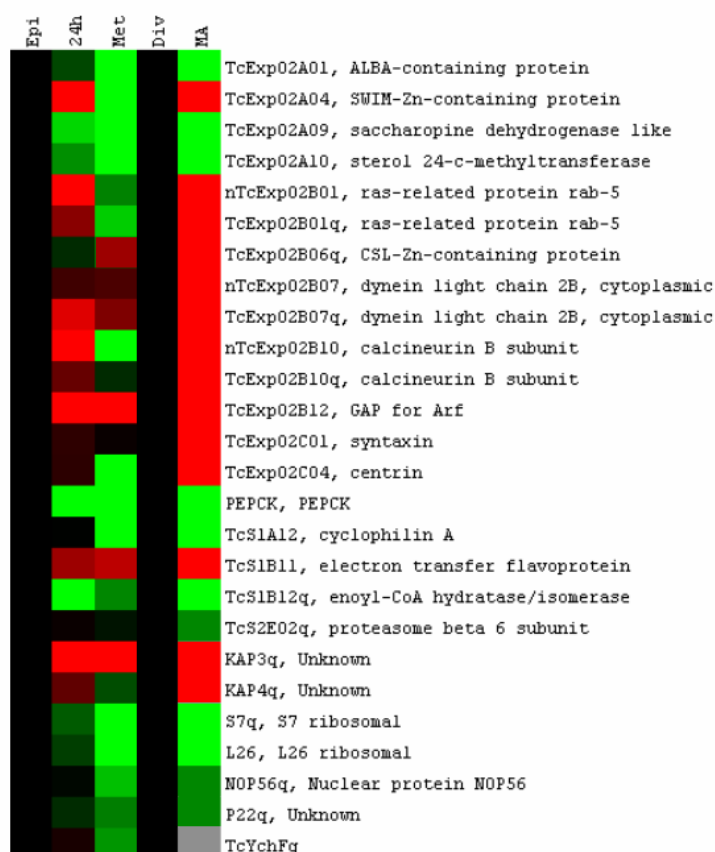


Figura 5.92. Gráfico *heatmap* mostrando a correlação entre transcriptoma e proteoma.

Em cada linha, proteínas estudadas; coluna – Epi, 24h e Met: extratos protéicos das fases epimastigota, intermediária da diferenciação (24 horas) e tripomastigotas metacíclicos; MA – consolidação dos dados de microarranjo em uma tendência geral. A coloração indica o sentido da modulação, sendo preto, ausência; vermelho, aumento; verde – diminuição; cinza – dado ausente. A graduação da cor indica o valor proporcional de modulação, para os extratos protéicos (tons mais escuros, modulação menor; tons mais claros, maior) e a tendência geral, para microarranjo (tom mais escuro, tendência moderada; mais claro, tendência forte). As letras q e n no nome do elemento representam, respectivamente, quantificação por quimioluminescência e dados normalizados para diminuir a variância.

De maneira geral, das 23 proteínas quantificadas, somente uma não apresentou padrão de modulação durante a metaciclogênese, a subunidade beta 6 do proteassomo (TcS2E02). Isso se deve provavelmente ao fato de que o sinal obtido para esse anticorpo foi muito fraco, mesmo com o uso de quimioluminescência e, portanto, esse resultado foi considerado não informativo.

Das 21 proteínas quantificadas e com dados do microarranjo, obteve-se resultados muito correlatos, cujo grau de concordância depende da forma como é classificado o padrão geral de metaciclogênese. Deve-se salientar a melhor qualidade do resultado obtido por quimioluminescência em relação ao método colorimétrico, o que pode ser visto para as proteínas identificadas por ambos os métodos (TcExp02B01, TcExp02B07 e TcExp02B10). Esse resultado é esperado, devido à maior sensibilidade do método quimioluminescente e pelo

fato de que as proteínas analisadas por ambos os métodos foram justamente aquelas que apresentaram um sinal muito fraco no método colorimétrico, sendo submetidas subsequente à quantificação por quimioluminescência.

Na análise mais estridente, 16 proteínas (76,2%) apresentaram um padrão coerente entre proteoma e transcriptoma (TcExp02A01, TcExp02A09, TcExp02A10, TcExp02B07, TcExp02B10, TcExp02B12, TcExp02C01, PEPCK, TcS1A12, TcS1B11, TcS1B12, Kap3, S7, L26, NOP56 e p22). Os critérios utilizados para a análise mais estridente foram a ocorrência de modulação dos níveis protéicos em 24 horas e/ou metacíclico, sendo que nos casos no qual a modulação foi vista para somente uma das amostras, deveria haver ausência total de modulação na outra.

Avaliou-se também a correlação entre transcriptoma e proteoma com um critério menos estridente, no qual o perfil de expressão diferencial em metacíclico, quando divergente do perfil de 24 horas aderido, não seria levado em consideração. De maneira geral, esse critério menos estridente também pode ser definido como a comparação da correlação entre transcriptoma e proteoma somente para 24 horas aderido, uma fase da diferenciação no qual a tradução é mais ativa.

Na análise menos estridente, todas as proteínas apresentaram um padrão de correlação geral entre transcriptoma e proteoma. As proteínas que foram incluídas no critério menos estridente são TcExp02A04 (padrão de aumento geral no microarranjo de DNA, forte aumento em 24 horas, forte diminuição em metacíclicos), TcExp02B01 (padrão de aumento geral no microarranjo de DNA, forte aumento em 24 horas, moderada diminuição em metacíclicos), TcExp02B06 (padrão de aumento geral no microarranjo, fraca diminuição em 24 horas, moderado aumento em metacíclicos), TcExp02C04 (padrão de aumento no microarranjo de DNA, aumento moderado em 24 horas, forte diminuição em metacíclicos) e Kap3 (padrão de aumento geral no microarranjo de DNA, forte aumento em 24 horas, moderada diminuição em metacíclicos).

6 Discussão

O presente trabalho teve como objetivo principal a melhoria na caracterização funcional de proteínas hipotéticas de *T. cruzi* que foram identificadas pelo uso de microarranjo de DNA como moduladas durante um processo muito interessante de diferenciação celular, a metaciclogênese, com ênfase em proteínas cuja expressão, medida a nível de mRNA polissomal, estivesse relativamente aumentada durante o processo.

Devido à complexidade dos fenômenos biológicos, com seus elementos e dinâmicas próprias, a genômica funcional não deve ser encarada como um resultado final, mas sim como um meio de selecionar novos genes para caracterização e também elucidar aspectos destes processos de difícil compreensão.

Com o crescente aumento do número de genomas completos seqüenciados, um grande desafio para os pesquisadores é obter a compreensão da expressão, função e regulação de um conjunto amplo, se não o total, de proteínas expressas no organismo. Através disto, será possível melhorar a compreensão de como os processos biológicos ocorrem ao nível molecular, como variam nos diferentes tipos de células, organismos, bem como nos estados alterados como, por exemplo, doenças (PATTERSON & AEBERSOLD, 2003; ZHU & *et al*, 2003). Sabe-se hoje que um gene não necessariamente codifica apenas uma proteína: o controle pós-transcricional como o *splicing* alternativo, poliadenilação, editoração do mRNA, entre outros; bem como o controle pós-traducional, dos quais são conhecidos aproximadamente 200 tipos diferentes, e ainda a proteólise e a compartimentalização, são desafios para a compreensão dos sistemas (PANDEY & MANN, 2000; GRAVES & HAYSTEAD, 2002). Visando isto, uma ferramenta que pode auxiliar e ser usada para se alcançar parte dos objetivos dessa nova era, é o microarranjos de DNA.

Desde 2001, iniciamos o projeto de análise por genômica funcional de *T. cruzi* por microarranjo, tendo recentemente incorporado abordagens proteômicas a esse arcabouço analítico, o que permitiu a definição de um conjunto muito grande de genes diferencialmente expressos, cuja implicação funcional aos processos biológicos estudados é relativamente desconhecida. Devido ao caráter de análise em larga escala que obtemos com o estudo da metaciclogênese por microarranjo de DNA, é necessária uma mudança na definição de estratégias apropriadas para responder às perguntas criadas pelo estudo genômico funcional. Portanto, é necessária a obtenção de dados de diversos elementos (genes, mRNA, proteínas) relativos, de certa forma, mais ao organismo como um todo, do que a certas definições

simplificadas e arbitrárias de vias biológicas selecionadas, as quais geralmente são abordadas pelo estudo de um ou pouco genes selecionados, a partir de conhecimento prévio já estabelecido.

Nesse sentido, é essencial o estabelecimento de novas metodologias de análise e de novas abordagens que permitam preencher as lacunas de conhecimento que temos, a fim de que possamos tirar maior proveito dos dados que são gerados de forma exponencial e contínua. Isso é especialmente importante no estudo da genômica funcional de *T. cruzi*, devido às peculiaridades de sua biologia molecular, principalmente no tocante à regulação da expressão gênica, bem como em razão da distância evolutiva grande em relação a outros organismos bem estudados, o que dificulta o estabelecimento empírico de atribuições funcionais pela simples similaridade de sequência aminoacídica primária.

Derivando o conhecimento adquirido com o uso do microarranjo de DNA, iniciou-se no ano de 2004 um projeto cujo principal objetivo é a determinação da estrutura tridimensional de proteínas moduladas durante os processos biológicos de *T. cruzi* estudados, com ênfase na metaciclogênese, o que permitirá a atribuição de função protéica não mais pela comparação de sua estrutura primária, mas sim pela comparação da estrutura terciária, a qual é muito mais informativa. Dentro do projeto de genômica estrutural de *T. cruzi*, optou-se pela criação dos clones contendo os genes de interesse a serem estudados em vetores da plataforma Gateway[®], o que permite uma maior automação do processo de clonagem e uma grande flexibilidade na utilização posterior desses clones em outros sistemas de avaliação funcional, como vetores de localização celular, de super-expressão induzida, de expressão em sistema homólogo e de identificação de complexos de interação protéica, entre outros.

Desta forma, iniciou-se um segundo projeto derivado, o qual visa à criação de um conjunto de vetores, dentro da plataforma Gateway[®], que permitirão a avaliação dos genes selecionados de acordo com os aspectos citados anteriormente. A seleção dos genes de *T. cruzi* que estão sendo incorporados é baseada nos dados de expressão diferencial obtidos pelo uso dos microarranjos e, atualmente, a base de dados existente no IBMP relativa ao transcriptoma de *T. cruzi* é ampla, se estendendo além do processo de metaciclogênese.

Dentro desse contexto, a presente dissertação de mestrado se insere, constituindo-se na primeira iniciativa de se aumentar o conhecimento funcional de uma parcela significativa de genes diferencialmente expressos, identificados por microarranjo. Por consistir em uma abordagem preliminar, a quantidade de genes estudados é relativamente grande quando comparada a estudos usuais de biologia molecular, mas representa uma pequena parcela dos

genes modulados em um único processo de diferenciação. Por exemplo, os 28 genes selecionados para esse trabalho constituem somente 2% dos aproximadamente 1200 genes modulados com um grau aceitável de confiança durante o processo de metaciclogênese e somente 0,3% dos aproximadamente 8000 genes distintos que estão presentes no genoma de *T. cruzi*.

Portanto, existe uma defasagem entre nossa capacidade atual de preencher as lacunas de informação funcional dos genes de *T. cruzi* e as demandas já existentes, criadas pelo uso de ferramentas de análise em larga escala. Nesse sentido, estamos estabelecendo um terceiro projeto amplo, denominado ORFeoma de *T. cruzi*, também inserido no projeto mais abrangente do estudo da genômica funcional de *T. cruzi*. O objetivo principal desse projeto é clonar todos os genes deste parasita em vetores de entrada da plataforma Gateway[®], possibilitando a avaliação de suas características funcionais, como localização celular, redes de interação protéica e redes de regulação gênica, dentro de uma perspectiva em larga escala.

A presente dissertação de mestrado também se insere dentro desse projeto de ORFeoma, por representar uma abordagem piloto da importância e viabilidade desse projeto maior, com relação à potencialidade de gerar informações biológicas novas e interessantes.

Nesse sentido, optou-se pela seleção de genes diferencialmente expressos por microarranjo durante a metaciclogênese, com ênfase nos que tivessem sua expressão relativa aumentada durante o referido processo. Além disso, selecionou-se como genes candidatos principais aqueles que estivessem anotados como proteínas hipotéticas conservadas com domínio funcional identificado. A justificativa para tal seleção é feita da seguinte forma:

- **Genes diferencialmente expressos:** estariam sendo regulados no processo da metaciclogênese e, portanto, teriam uma maior probabilidade de serem agentes diretos das mudanças que ocorrem no mesmo;
- **Expressão aumentada:** de maneira geral, os genes que apresentam sua expressão diminuída durante a metaciclogênese estão enriquecidos em proteínas com função conhecidas e participantes de vias bioquímicas clássicas (PROBST, 2005), sendo que o mesmo não ocorre para os genes com expressão relativa aumentada. Além disso, genes cujo produto estivesse sendo requerido no processo de diferenciação constituem um grupo de maior interesse;
- **Proteínas hipotéticas:** esses genes codificam para proteínas cuja atribuição funcional é inexistente ou muito genérica, necessitando, portanto, de um

aprofundamento de suas características funcionais para que hipóteses sejam feitas a respeito do que ocorre na metaciclogênese;

- **Conservação evolutiva:** aumentam a confiabilidade de que a proteína estudada seja realmente encontrada no organismo;
- **Domínio funcional:** permitem a definição de diretrizes que auxiliam no direcionamento do trabalho, uma necessidade vital para o desenvolvimento de uma dissertação de mestrado em tempo hábil.

Visando conciliar duas necessidades divergentes, quer seja: a demanda pela geração de uma grande quantidade de informações funcionais relativas aos genes diferencialmente expressos durante a metaciclogênese; e a necessidade de obter resultados novos e interessantes e claros dentro da perspectiva temporal de conclusão de uma dissertação de mestrado, optou-se pela idealização lógica de uma linha de trabalho que avaliasse um conjunto relativamente amplo de genes e a definição de salvaguardas necessárias para a obtenção de resultados.

Dessa forma, definiu-se arbitrariamente que o conjunto primário de genes a serem caracterizados funcionalmente seria em número de cinco, os quais foram denominados de genes principais. Para evitar que falhas no fluxograma do presente trabalho, ilustrado na Figura 4.2, impedissem a obtenção desse objetivo, criaram-se dois níveis de redundância: a inclusão de dois genes principais adicionais e a seleção de outros sete genes com características semelhantes aos genes principais, e que foram denominados de genes secundários. Com isso, aumentou-se a probabilidade de que, ao final do trabalho, dados funcionais fossem obtidos para pelo menos cinco genes codificadores de proteínas hipotéticas conservadas, contendo domínio funcional caracterizado e diferencialmente expressos durante a metaciclogênese.

A fim de estudar a coerência funcional dos padrões de co-expressão em *T. cruzi*, optou-se por selecionar sete grupos com padrões distintos de expressão diferencial, e dentro de cada um desses grupos, selecionou-se um gene principal e secundário. A partir dessa definição metodológica, foi evidenciado que seria interessante acrescentar, dentro de cada grupo de co-expressão, um terceiro gene, cujo conhecimento funcional fosse maior, isto é, cujas características conhecidas fossem distintas das compartilhadas pelos genes principal e secundário, o qual foi denominado de gene guia.

Como um dos objetivos de caracterização funcional do presente trabalho é a localização celular dos genes selecionados, a inclusão de um gene cujo conhecimento funcional fosse maior, incluindo a sua localização celular, possibilitaria a obtenção de resultados esperados previamente, aumentando a confiabilidade dos resultados obtidos para as outras proteínas menos caracterizadas.

Finalmente, incluiu-se em cada grupo um gene, denominado acessório, o qual compartilha as características gerais dos genes principais e secundários, portanto aumentando ainda mais a redundância. No entanto, o gene acessório possui duas características específicas que diminuem a probabilidade de sucesso na sua caracterização funcional, qual seja a inexistência de ortólogos em outros organismos e a ausência de domínios funcionais caracterizados. Embora esses dois fatores possam indicar um possível erro de predição e uma menor quantidade de informações prévias, eles também reforçam a possibilidade de que, ao se caracterizar funcionalmente os mesmos, estejamos agregando informações mais interessantes, inovadoras e diretamente relacionadas às peculiaridades de *T. cruzi*.

Dentro desse arcabouço lógico, definimos alguns objetivos práticos que deveriam ser alcançados com relação a todo esse conjunto de genes:

- A obtenção de clones de entrada na plataforma Gateway[®] para todos os genes selecionados;
- Expressão de proteína e produção de anti-soro para pelo menos 14 proteínas (metade do conjunto total), sendo pelo menos sete proteínas hipotéticas (dos tipos principal, secundário ou acessório).
- Imunolocalização de pelo menos cinco proteínas hipotéticas (dos tipos principal, secundário ou acessório);
- Obtenção de *western blot* quantitativo, durante a metaciclogênese, para pelo menos 10 proteínas distintas, independentemente de seu tipo;
- Análise funcional detalhada de pelo menos uma proteína hipotética (dos tipos principal, secundário ou acessório).

Com a definição desses objetivos práticos e com a inclusão de diversos níveis de redundância na seleção dos genes candidatos seguiu-se uma abordagem de avanço seqüencial, unidirecional, em relação ao fluxograma estabelecido. Isso se reflete na definição de somente um protocolo para cada uma das etapas, com pequenos ajustes de acordo com a necessidade, desde que não atrasassem a conclusão da etapa atual; e o prosseguimento para a etapa

seguinte com os elementos que fossem obtidos na etapa atual, sem que se realizasse uma nova tentativa de obtenção.

Dos 28 genes iniciais, obteve-se a amplificação, clonagem em vetor de entrada (pDONR) e passagem para vetor de destino para expressão protéica em *E. coli* (pDEST) para todos os genes selecionados. Após a realização do seqüenciamento visando certificar a clonagem correta do produto amplificado, evidenciou-se que para o clone TcExp02B03 havia uma deleção de um nucleotídeo no iniciador *forward*, criando uma mudança no quadro de leitura, o que levou à exclusão desse gene das etapas posteriores.

É possível evidenciar a alta eficiência do sistema de clonagem selecionado para o presente trabalho, o que permite concluir que essa fase, na construção do ORFeoma completo de *T. cruzi*, que se baseará na mesma plataforma, não consistirá em um ponto problemático dentro da linha de produção de clones.

Após a obtenção dos clones de expressão, procedeu-se à expressão e purificação protéica dos 27 clones restantes, etapa esta que representou um grande gargalo para o prosseguimento da análise. Foi possível obter 19 proteínas expressas (70,4%), todas insolúveis. Como o principal interesse da expressão protéica para o presente trabalho é a obtenção de anticorpos policlonais para a realização de *western blot* quantitativo e imunolocalização, a insolubilidade das proteínas não representou um resultado adverso e a expressão de todas as 19 proteínas foi considerada de qualidade suficiente para o prosseguimento da análise. Deve-se salientar que foi utilizado somente um único protocolo de expressão, dentro da lógica de trabalho adotada e descrita acima, e que é muito provável que a utilização de outros protocolos possa, futuramente, aumentar a quantidade de proteínas expressas e sua solubilidade.

Quanto ao processo de purificação das mesmas, foram utilizados três protocolos distintos conforme descrito no item 4.11. Isso ocorreu devido às características individuais das proteínas que podem responder melhor a diferentes condições de purificação. Primeiramente, a purificação utilizando colunas baseadas na ligação da cauda de histidina ao níquel foi feita por duas abordagens diferentes: purificação por eluição utilizando concentrações crescentes de imidazol e por decréscimo do pH.

Mesmo após a realização desse processo de purificação por afinidade, foi possível evidenciar proteínas contaminantes em pequena quantidade. Para aumentar o grau de pureza das proteínas estudadas, optou-se pela realização de uma purificação por SDS-PAGE. No final, das 19 proteínas obtidas na etapa de expressão, conseguiu-se 18 proteínas purificadas (94,7% do conjunto purificado; 64,3% do conjunto total), tendo-se eliminado do restante do

processamento a proteína TcExp02A07, a qual já apresentava uma baixa expressão e devido a perdas cumulativas durante as etapas de purificação.

Na etapa de obtenção de anticorpos policlonais, foram usados dois animais para cada proteína, totalizando 36 camundongos, os quais foram inoculados em dois locais diferentes, respectivamente pelas vias subcutânea e intra-peritoneal. Destes animais, três morreram no decorrer do processo, sendo que dois haviam sido inoculados com a mesma proteína (TcExp02B04) e, portanto, não foi possível o prosseguimento desta para as etapas subseqüentes. Ao final, foram obtidos 33 soros (91,6% do total de animais inoculados), referentes a 17 proteínas (60,7% do conjunto inicial). Dos 36 soros obtidos antes das inoculações, denominados pré-imune, nenhum apresentou reação contra o extrato de *T. cruzi*, permitindo assim a continuidade do processo, sem descarte ou troca de animal.

Visando a validação e verificação da qualidade dos soros, foram realizados ensaios de *western blot*, divididos em duas etapas. Primeiramente, foram utilizados a proteína recombinante e o extrato da forma epimastigota de *T. cruzi*, para verificar a qualidade do anti-soro obtido. Na segunda etapa, foram utilizados extratos das formas epimastigota, fase intermediária da diferenciação (aderido por 24 horas) e tripomastigota metacíclico, para avaliar a expressão diferencial das proteínas durante a metaciclogênese. Nessa etapa foi possível identificar os anti-soros com padrão de reação com qualidade aceitável e, após a consolidação dos soros duplicados para uma mesma proteína, foram selecionados onze soros (39,3% do conjunto total, 64,7% dos anti-soros produzidos), sendo este o segundo maior obstáculo influenciando a produtividade do processo geral.

Conforme mencionado acima, os onze anti-soros foram empregados na análise por *western blot* do padrão de expressão diferencial durante a metaciclogênese, associado a doze anti-soros produzidos em outros projetos do IBMP, totalizando 23 proteínas diferencialmente expressas durante a metaciclogênese.

Na literatura, existem diversos trabalhos utilizando microarranjos para avaliar a expressão diferencial de tripanossomatídeos (DIEHL *et al.*, 2002; GUIMOND *et al.*, 2003; MINNING *et al.*, 2003; SAXENA *et al.*, 2003; AKOPYANTS *et al.*, 2004; ALMEIDA *et al.*, 2004; BAPTISTA *et al.*, 2004; BREMS *et al.*, 2005; BAPTISTA *et al.*, 2006; HOLZER *et al.*, 2006; McNICOLL *et al.*, 2006; LEIFSO *et al.*, 2007; SAXENA *et al.*, 2007; SRIVIDYA *et al.*, 2007). Esses trabalhos avaliam uma ampla gama de condições biológicas de diversos organismos, incluindo o transcriptoma de *T. cruzi* (MINNING *et al.*, 2003; BAPTISTA *et al.*, 2004; BAPTISTA *et al.*, 2007). Uma característica geral desses trabalhos é a identificação de

um número muito pequeno de genes diferencialmente expressos, geralmente entre 0,5% e 5% do total presente no microarranjo, mesmo quando comparando formas evolutivas funcional e morfológicamente distintas entre si.

Esses resultados são diametralmente opostos aos obtidos por nosso grupo e, portanto, cria-se um paradoxo cuja elucidação é crucial para a determinação da validade do estudo do transcriptoma de tripanossomatídeos como uma ferramenta de análise biológica. Existem diversas diferenças técnicas entre os trabalhos supracitados e os desenvolvidos no IBMP, os quais, em conjunto, explicam em grande parte a diferença observada. Estes são, resumidamente, o uso de mRNA polissomal, a amplificação do mRNA, a marcação da segunda fita do cDNA, o uso de microarranjos baseados em ESTs e a maior quantidade de replicatas de hibridação (PROBST, 2005).

No entanto, é necessária a comprovação dos resultados obtidos por nosso grupo, devido ao seu impacto, o que foi feito utilizando-se quatro abordagens principais (PROBST *et al.*, em preparação):

- **Análise do transcriptoma por PCR em tempo real quantitativa (qPCR):** corrobora os resultados do microarranjo. Porém, avalia a mesma representação de expressão diferencial, o transcriptoma.
- **Coerência funcional das modulações observadas:** através da análise estatística usando anotação funcional dos bancos de dados KEGG (Kyoto Encyclopaedia of Genes and Genomes) e GO (Gene Ontology), evidenciar que as modulações observadas são funcionalmente coerentes, o que evidencia que as mesmas são importantes do ponto de vista funcional.
- **Proteômica diferencial em larga escala:** usando dados obtidos da análise proteômica em larga escala, identificar proteínas diferencialmente expressas e verificar se o mesmo padrão é observado nos dados do microarranjo.
- **Proteômica diferencial em baixa escala:** através da produção de anti-soros, evidenciar padrões de expressão protéica diferencial por *western blot*.

Com os resultados obtidos para essas quatro abordagens diferentes, foi possível obter um padrão geral de grande correlação entre as modificações transcriptômicas e proteômicas, o que evidencia a qualidade dos resultados gerados pela análise de transcriptoma por microarranjo realizada no IBMP. A análise realizada no presente trabalho, a quantificação por

western blot para 23 proteínas diferencialmente expressas, está dentro da metodologia da quarta abordagem supracitada.

Conforme mencionado anteriormente, obteve-se um alto grau de concordância entre os resultados de microarranjo e *western blot*, quando os dados quantitativos obtidos foram analisados em relação a padrões gerais de aumento ou diminuição durante a metaciclogênese. Segundo os critérios de alta e baixa estringência, respectivamente 76,2% e 95,2% dos genes/proteínas estudados mostraram concordância, um resultado muito contundente e, de certa forma, inesperado.

A importância da modulação do transcriptoma como elemento chave na modificação do perfil protéico e, conseqüentemente, das características funcionais de uma célula, é um fato basal na utilidade do uso de microarranjos para o estudo de processos biológicos. No entanto, a comparação dos dados de transcriptoma e proteoma geralmente apresenta uma correlação relativamente baixa (FUTCHER *et al.*, 1999; IDEKER *et al.*, 2001; GRIFFIN *et al.*, 2002; WASHBURN *et al.*, 2003; BASLER *et al.*, 2006; NISSOM *et al.*, 2006; SCHERL *et al.*, 2006), embora alguns estudos mostrem o contrário (GYGI *et al.*, 1999; UNWIN *et al.*, 2005; MIJALSKI *et al.*, 2005).

Isso não é de todo inesperado, pois os níveis protéicos são ditados somente em parte pelos níveis de mRNA, sendo que diversas outras fontes de controle, como degradação protéica seletiva, podem modificar os níveis protéicos. Além disso, a quantificação por proteômica e microarranjo são métodos relativamente ruidosos, sendo que até mesmo diferentes plataformas comerciais de microarranjo apresentam correlação média entre os resultados (JÄRVINEN *et al.*, 2004), embora esses resultados possam variar de acordo com os critérios utilizados (PETERSEN *et al.*, 2005).

Independentemente do grau de concordância existente entre transcriptoma e proteoma, o resultado principal da análise feita nesse trabalho é a determinação de que as diferenças identificadas no microarranjo de *T. cruzi*, as quais são em grande número e destoando de forma marcante dos demais estudos de transcriptoma de tripanossomatídeos, influem nos níveis protéicos e, portanto, nas modulações que o *T. cruzi* tem que sofrer para alterar sua composição protéica em resposta aos estímulos ambientais, externos ou internos.

Embora isso possa parecer óbvio à primeira vista, diversas observações podiam indicar o contrário, as quais são:

- Diversos trabalhos demonstrando a ausência de modulação dos níveis de mRNA em tripanossomatídeos, tendo sido inclusive sugerido que não haja regulação significativa das proteínas nesses organismos (LEIFSO *et al.*, 2007);
- As características peculiares de tripomastigotas metacíclicos, uma forma de resistência cuja principal atribuição evolutiva é a infecção do hospedeiro mamífero. Isso se reflete na ausência de replicação do DNA e menor taxa de transcrição (ELIAS *et al.*, 2001; ELIAS *et al.*, 2002) e tradução (GOLDENBERG *et al.*, 1985) dessa fase, o que indicaria uma importância menor da quantificação do mRNA em metacíclicos, mas não durante a metaciclogênese. Reforçando essa menor influência dos níveis de mRNA mensageiro na produção protéica de metacíclicos está a descrição de um mRNA que está associado a polissomos na metaciclogênese, mas cuja proteína se encontra ausente (NARDELLI *et al.*, 2007) e a verificação, nos próprios dados de microarranjo da metaciclogênese de *T. cruzi*, de que genes envolvidos na replicação de DNA e transcrição de RNA estariam aumentados em metacíclicos (PROBST, 2005).

Com os resultados obtidos no presente trabalho, foi possível evidenciar que as modificações que ocorrem no mRNA polissomal tem repercussão clara no proteoma de *T. cruzi*, reforçando a utilidade da técnica de microarranjo no estudo de questões biológicas de *T. cruzi*, bem como comprovando a confiabilidade do método de análise utilizado no projeto de genômica funcional do IBMP. Esse resultado representa uma importante contribuição para o melhor entendimento da biologia molecular de *T. cruzi*.

Com relação à alta concordância evidenciada entre transcriptoma e proteoma, diversos fatores podem ser levantados para explicar esse resultado, os quais são:

- Utilização de mRNA polissomal, mais correlacionado com os níveis protéicos;
- Seleção de genes candidatos baseados em critérios mais intensos de confiabilidade e expressão diferencial;
- Utilização de um método mais sensível e preciso de quantificação protéica, imunoquantificação por *western blot*, ao invés de uma análise proteômica em larga escala (MudPIT ou 2D-PAGE);
- Avaliação de um processo temporal, permitindo a compensação por defasagens da repercussão de uma modificação nos níveis do mRNA e proteína.

Esse último ponto é especialmente relevante na comparação de baixa estringência entre o transcriptoma e o proteoma. De maneira geral, a maioria dos eventos de modulação do mRNA ocorre em uma fase relativamente inicial da metaciclogênese (ver Figura 4.4 a Figura 4.10), permitindo a modificação dos níveis protéicos em parasitas em diferenciação por 24 horas; no entanto, essa modulação do nível de mRNA poderia ter um impacto menor em tripomastigota metacíclico, o qual apresenta uma diminuição da tradução (GOLDENBERG *et al.*, 1985; PROBST, 2005) e, conseqüentemente, os níveis protéicos existentes nessa fase do parasita tenderiam a ser definidos mais pela meia-vida das proteínas, de acordo com a sua cinética de degradação, do que pelos níveis de mRNA polissomal. Dessa forma, a análise com baixa estringência evita a interferência de situações nas quais a meia-vida curta da proteína daria o resultado final da quantificação por *western blot* em metacíclicos (no presente trabalho, quatro das 21 proteínas analisadas ou 19% do total).

Outra importante contribuição do presente trabalho foi o aumento de informação obtido para as proteínas estudadas, com relação a aspectos bioinformáticos de anotação e de imunolocalização.

Com relação às proteínas do grupo 1, foi possível evidenciar que:

- TcExp02A01, previamente identificada como contendo um domínio relacionado à interação com DNA nuclear ou RNA citoplasmático, é na verdade citoplasmática, pelos dados de imunofluorescência obtidos. Além disso, identificou-se um domínio RGG-box, presente somente em proteínas de protozoários, o qual também interage com RNA, reforçando a importância dessa proteína como sendo de ligação ao RNA;
- TcExp02A08, inicialmente anotada como proteína hipotética, teve um ortólogo identificado em *L. infantum*, e deve ser, portanto, renomeada para proteína hipotética conservada;
- Anotação do domínio funcional sacaropina-desidrogenase para TcExp02A09 e sua identificação como uma proteína citoplasmática relacionada provavelmente à biossíntese de lisina;
- Anotação dos domínios funcionais metil-transferase e esterol-metiltransferase para TcExp02A10, indicando sua ligação à uma família de metil-transferases associadas a S-adenosil-metionina (SAM) e a determinação de sua localização citoplasmática por imunofluorescência.

Com relação às proteínas do grupo 2, foi possível evidenciar que:

- TcExp02A02 é uma proteína-quinase e contém uma grande região protéica sem domínios preditos. As proteínas de eucariotos superiores mais similares a ela estão anotadas como quinase 2 do fator de iniciação da tradução (eIF2AK2). Em geral, eIF2AK2 é composto somente pelo domínio quinase e, portanto, caso TcExp02A02 seja realmente o ortólogo dessa proteína, possui uma região de ~600 aminoácidos a qual potencialmente contém alguma estrutura funcional acessória.
- TcExp02A12 apresentou resíduos extremamente conservados de uma zinc finger, porém não foi identificado um domínio claro nessa proteína. Portanto, essa proteína é provavelmente uma zinc finger contendo um novo domínio ainda não caracterizado.
- TcExp02B01 está anotada como rab5 e foi feita sua localização em uma região entre o cinetoplasto e o flagelo.

Com relação às proteínas do grupo 3, foi possível evidenciar que:

- TcExp02A03 é uma ciclofilina do tipo FKBP, contendo um domínio degenerado;
- TcExp02B02, inicialmente anotada como proteína hipotética, teve um ortólogo identificado em *L. infantum*, e deve ser, portanto, renomeada para proteína hipotética conservada.

Com relação às proteínas do grupo 4, foi possível evidenciar que:

- TcExp02B06 é uma proteína relacionada à via de produção da diftamida, uma modificação pós-traducional de uma histidina, encontrada somente em eEF2. A via de síntese da diftamida é composta por cinco genes, os quais foram totalmente identificados em *T. cruzi*. TcExp02B06 é DPH3 e contém o domínio CSL, DPH4 (5788.t00002) contém um domínio CSL degenerado em *T. cruzi*, associado a um domínio DnaJ, DPH1 (3750.t00001), DPH2 (4935.t00029) e DPH5 (10960.t00003).
- TcExp02B07 foi identificada como sendo uma proteína localizada preferencialmente no flagelo.

Com relação às proteínas do grupo 5, foi possível evidenciar que:

- TcExp02B08, inicialmente anotada como proteína hipotética, teve ortólogos identificados em *T. congolense* e *T. vivax*, e deve ser, portanto, renomeada para proteína hipotética conservada;

- TcExp02B10 foi identificada como sendo uma proteína localizada preferencialmente no flagelo.

Com relação às proteínas do grupo 6, foi possível evidenciar que:

- TcExp02B11, anotada inicialmente como uma proteína hipotética, realmente não apresenta ortólogos em tripanossomatídeos. No entanto, uma análise mais detalhada a identificou como membro de uma família de proteínas relacionadas, composta por quatro grupos protéicos, sendo dois compartilhados em tripanossomatídeos e dois restritos a *T. cruzi*, incluindo TcExp02B11. Essa família protéica tem como características principais o fato de serem proteínas de membrana sem nenhum domínio funcional identificado;
- TcExp02B12 foi localizada em uma região entre o flagelo e o cinetoplasto;
- TcExp02C01 foi localizada em uma região entre o flagelo e o cinetoplasto.

Com relação às proteínas do grupo 7, foi possível evidenciar que:

- TcExp02A07 é possivelmente uma MAP quinase;
- TcExp02C04 foi localizada preferencialmente no cinetoplasto.

As proteínas com imunofluorescência foram selecionadas para um maior detalhamento, devido à importância da imunolocalização para a caracterização funcional. Dos onze soros obtidos, dois foram eliminados devido à reação cruzada do soro pré-imune, observada somente na imunofluorescência controle, pois todos soros pré-ímmes não reagiram com um extrato de *T. cruzi* na etapa inicial de seleção. Dos nove soros restantes, sete apresentaram um padrão de localização bem definido, um soro produziu um padrão de dispersão no citoplasma, e outro soro obteve um padrão mais intenso no cinetoplasto e um padrão disperso pelo citoplasma.

A imunolocalização da proteína TcExp02A01 apresentou um padrão granular disperso pelo citoplasma do parasita. O domínio ALBA, presente nesta proteína, está envolvido na ligação a ácidos nucléicos. Dentro desta função primária, é descrita na literatura a capacidade de atuar na compactação de DNA em arqueobactérias (BELL *et al.*, 2002; MARSH *et al.*, 2005; SANDMAN & REEVE, 2005) e de estar associada à ligação com RNA com função de estabilização e/ou ligada ao metabolismo de RNA (ARAVIND *et al.*, 2003; GUO *et al.*, 2003; MARSH *et al.*, 2005).

Segundo ARAVIND *et al.* (2003), o domínio ALBA faz parte de uma superfamília com função básica de ligação a ácidos nucleicos, sendo que em eucariotos a função está baseada na similaridade entre algumas proteínas de função conhecida como a IF3-C (Fator de iniciação da tradução 3) e YhbY (Domínio de ligação a RNA presente em bactérias e plantas). Outra diferença é o tamanho das proteínas, pois em arqueobactérias as proteínas com domínio ALBA clássicas possuem em média 10 kDa, e em eucariotos possuem tamanhos variados, muitas vezes maiores, o que ocorre com TcExp02A01.

É interessante ressaltar que em *T. cruzi* há quatro proteínas minimamente distintas apresentando o domínio ALBA. TcExp02A01 na verdade é uma representação individual de dois parálogos, justapostos no genoma, que são muito similares em toda a sua extensão, diferindo somente nos últimos 10 aminoácidos de sua porção carboxi-terminal. Esses dois parálogos apresentam essas mesmas características nas outras espécies do gênero *Trypanosoma* seqüenciadas, mas somente um ortólogo em *Leishmania*.

As outras duas proteínas contendo o domínio ALBA apresentam um tamanho (13kDa) mais próximo ao das proteínas Alba clássicas de arqueobactérias, sendo compostas praticamente somente por esse domínio. Esses dois genes distintos também se encontram justapostos no genoma de todos os tripanossomatídeos seqüenciados, mas apresentam somente 41% de similaridade em sua seqüência aminoacídica.

Em TcExp02A01, existe uma região conservada na extremidade carboxi-terminal rica em argininas e glicinas, cujo motivo, denominado RGG-box, está presente em diversas proteínas ligadoras de RNA (McBRIDE *et al.*, 2005). O padrão granular distribuído pelo citoplasma, encontrado na imunolocalização de TcExp02A01, associado a esse motivo de ligação ao RNA, é compatível com a atividade relacionada ao metabolismo e/ou estabilização do RNA, pois não foi encontrado indícios de presença de sinal da proteína no núcleo ou cinetoplasto.

A fim de identificar os possíveis ortólogos de TcExp02A01 em outros organismos, foi feito uma procura bioinformática por proteínas contendo o domínio ALBA e pelo menos dois motivos RGG presentes em uma mesma região. Interessantemente, o domínio ALBA não foi encontrado em fungos (filamentosos ou leveduras) e somente em protozoários foram identificadas proteínas contendo tanto o domínio ALBA quanto o motivo RGG (gêneros *Trypanosoma*, *Leishmania* e *Plasmodium* e em *D. discoideum*), o que indica a restrição evolutiva desse tipo protéico.

Devido ao padrão granular identificado para ALBA no citoplasma de *T. cruzi* e ao fato de ter sido caracterizada como uma proteína de ligação ao RNA potencial, resolvemos avaliar se

essa proteína estaria associada a corpúsculos citoplasmáticos que ocorrem em *T. cruzi* preferencialmente em situações de carência nutricional.

HOLETZ *et al.* (2007) identificaram a presença de uma RNA helicase, TcDhh1, em focos citoplasmáticos os quais aumentavam em número em parasitas sobre estresse nutricional. A helicase Dhh1 é um marcador de uma estrutura citoplasmática denominada *P-bodies* encontrada em fungos e mamíferos, a qual está envolvida ou na degradação 5'-3' de mRNAs desadenilados e desencapados ou no armazenamento de mRNAs ainda tradutíveis. CASSOLA *et al.* (2007) aprofundaram a caracterização molecular desses corpúsculos citoplasmáticos e identificaram diversas outras proteínas como presentes, incluindo a própria TcDhh1, TcXRNA (a enzima envolvida na degradação 5'-3') e proteínas de ligação ao RNA (TcUBP1, TcUBP2, TcRBP3, TcRBP4, TcRBP5a, TcRBP6b, TcPABP1, TcPABP2, TceIF4E). Esses autores também evidenciaram a presença desses focos em parasitas sobre privação de fonte de carbono. Baseado na composição identificada desses corpúsculos, foi proposto que eles eram uma nova entidade de armazenamento de mRNA, pois eram compostos de proteínas presentes tanto em *p-bodies* quanto em grânulos de estresse (SG), duas estruturas distintas em eucariotos superiores. Possivelmente esses corpúsculos constituem uma estratégia de proteção transiente de transcritos em situações de privação de nutrientes (CASSOLA *et al.*, 2007).

Visando identificar se TcExp02A01 está associada aos corpúsculos citoplasmáticos identificados por HOLETZ *et al.* (2007) e CASSOLA *et al.* (2007), epimastigotas em fase estacionária de crescimento foram submetidos a estresse nutricional em meio TAU, segundo o protocolo de metaciclogênese *in vitro* (BONALDO *et al.*, 1988), e foi realizada uma imunofluorescência com anticorpos anti-TcDhh1 e anti-TcExp02A01, cujo resultado pode ser visto na Figura 6.1. É possível identificar claramente que as proteínas TcDhh1 e TcExp02A01 apresentam-se co-localizadas em corpúsculos citoplasmáticos distintos, o que é evidenciado em alguns parasitas. Interessantemente, em outros parasitas essas duas proteínas não apresentam co-localização, estando ainda presentes em corpúsculos citoplasmáticos distintos. Esse resultado sugere que os corpúsculos citoplasmáticos de *T. cruzi* são na verdade uma população heterogênea e que a dinâmica de tradução, estocagem e degradação dos mRNAs é complexa. Esses dados são preliminares e serão reanalisados futuramente. No entanto, é possível evidenciar que TcExp02A01 é uma proteína citoplasmática de ligação ao mRNA que se encontra, pelo menos parcialmente, associada às estruturas *p-bodies-like* de *T. cruzi*.

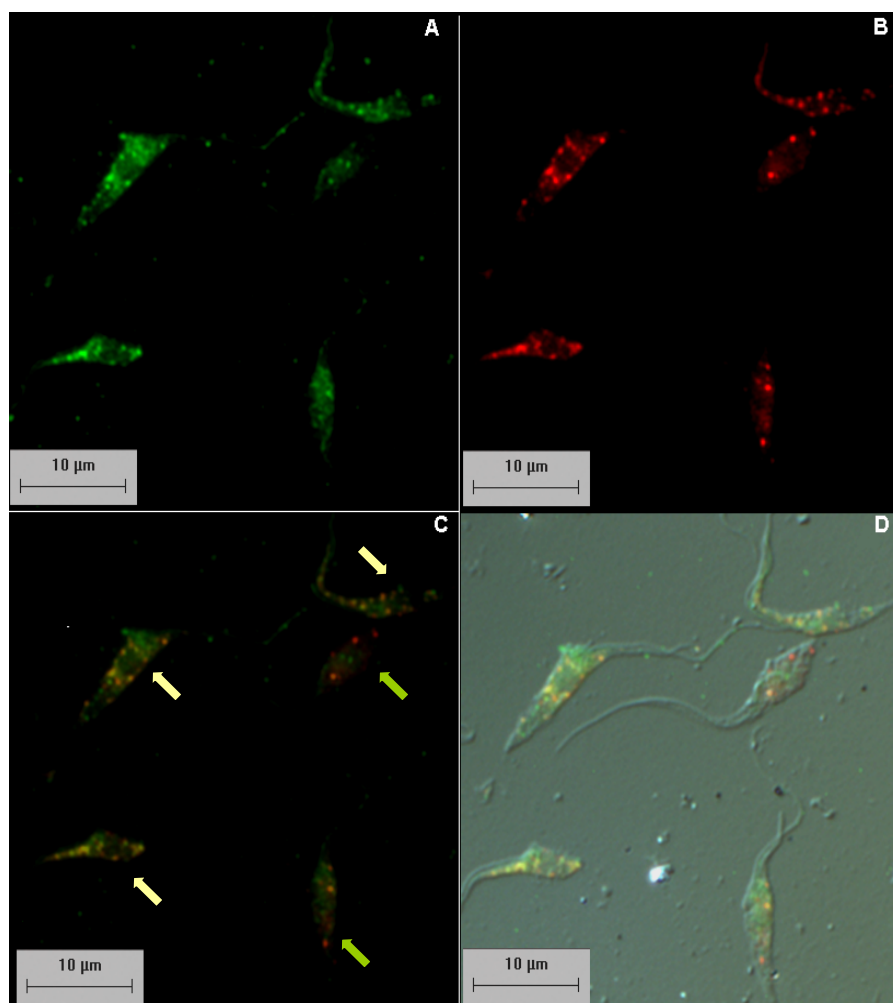


Figura 6.1. Imunolocalização de TcDhh1 e TcExp02A01.

Anti-TcDhh1 (verde, A), anti-TcExp02A01 (vermelho, B); C) sobreposição de A e B; D) contraste diferencial de fase sobreposto a C. Setas amarelas, parasitas apresentando co-localização de TcDhh1 e TcExp02A01; Setas verdes, parasitas mostrando corpúsculos citoplasmáticos contendo TcDhh1 e TcExp02A01 sem co-localização.

A definição de TcExp02A01 como uma proteína de ligação ao RNA reforça a necessidade de caracterizar quais são os seus alvos moleculares, para que possamos aprofundar nosso conhecimento sobre os mecanismos de regulação de expressão gênica pós-transcricional em *T. cruzi*. Nesse sentido é necessário avaliar quais seriam os mRNAs associados a TcExp02A01, além de identificar elementos presentes nas regiões 3'-UTR destes, responsáveis pela interação mRNA-proteína.

A abordagem a ser realizada consistirá em se fazer uma imunoprecipitação de TcExp02A01 em extrato protéico de *T. cruzi* e na identificação de mRNAs associados através do uso de microarranjos de DNA, em uma abordagem denominada ribonômica. Após a identificação desses alvos, será feita a procura de motivos sequenciais presentes nas regiões

3'-UTR e compartilhados pelos mRNAs co-precipitados. Finalmente, a comprovação da ligação entre TcExp02A01 e essas regiões do mRNA será feita através da técnica de EMSA (*electrophoretic mobility shift assay*), que consiste na separação eletroforética de complexos de proteína-mRNA, os quais são comparados com o padrão de mobilidade da sequência de RNA alvo. Caso haja interação entre a proteína analisada e a sequência alvo de RNA, esse complexo irá migrar diferentemente do RNA puro.

O gene codificador da proteína hipotética contendo o domínio sacaropina-desidrogenase (TcExp02A09), responsável pela etapa final da biossíntese de lisina, apresentou na imunolocalização um padrão mais intenso de fluorescência na região do cinetoplasto seguida de fluorescência dispersa pelo citoplasma. Este duplo padrão de localização encontrado pode ser devido à inespecificidade do soro, visto que os ensaios de *western blot* revelaram o reconhecimento de duas proteínas, uma com o tamanho esperado e outra com um tamanho menor conforme mostrado na Figura .

Essa proteína é extremamente conservada em diversos organismos. Na tentativa de identificar a proteína de tamanho inesperado, pretende-se a princípio realizar a verificação da qualidade da proteína recombinante a fim de verificar possível contaminação. Outra possibilidade é a separação das bandas por SDS-PAGE, seguida por identificação proteica por espectrometria de massa.

Outra forma de se obter o padrão de localização esperado de TcExp02A09 seria realizar a imunolocalização em tripomastigota metacíclico, pois no ensaio de *western blot* a proteína contaminante não foi identificada nessa fase.

Independentemente das análises a serem feitas, algumas observações indicam que essa proteína é possivelmente mitocondrial. Em primeiro lugar, a busca de similaridade por BLASTp identificou muitas proteínas de diversos organismos, desde cianobactérias até mamíferos, com uma região de similaridade muito semelhante, o que indica conservação funcional em um conjunto diverso de organismos evolutivamente distantes; além disso, a imunolocalização apresenta um sinal forte no cinetoplasto. BLEMINGS *et al.* (1994) sugere a ocorrência na mitocôndria de células do fígado de ratos, enquanto ZHU *et al.* (2000) sugere a ocorrência desta enzima na parte citoplasmática em *Arabidopsis thaliana*.

Com relação à imunolocalização do gene TcExp02A10 que codifica uma proteína hipotética conservada de 40,5 kDa, apresentando um domínio denominado de Sterol_MT_C (Sterol methyltransferase C-terminal), verifica-se um padrão citoplasmático, com uma maior emissão de fluorescência em torno da região nuclear. Este domínio está presente na região

carbóxi-terminal de metiltransferases, estando presente principalmente em fungos e plantas envolvidos com metilação de esteróis (BOUVIER-NAVE *et al.*, 1998). Em tripanossomatídeos, ergosterol e esteróis alquilados na posição 24 são componentes presentes em alta concentração nas membranas, uma situação semelhante a que ocorre em fungos e plantas. A biossíntese de ergosterol e esteróis que necessitam desalquilação na posição 24 da cadeia é um passo que não está presente na via do colesterol, presente nas células de mamíferos. Para que esta síntese ocorra é necessária a presença da proteína 24-esterol metiltransferase sendo esta a etapa chave na diferença da formação de ergosterol e colesterol (GROS *et al.*, 2006).

A imunolocalização da proteína referente ao gene TcExp02B01 apresentou um padrão bem definido, localizando-se entre a bolsa paraflagelar e o cinetoplasto, região esta normamente ocupada pelo Complexo de Golgi em *T. cruzi*. O domínio RAB é característico de proteínas com função GTPase que está envolvida no controle do tráfego intracelular, atuando como receptor da endocitose e na dinâmica endossomal. Dentro desta perspectiva, Rab5 regula o transporte de vesículas, mediado por clatrina, entre a membrana plasmática e os endossomos primários (recentes), além das fusões entre os endossomos primários (ZERIAL & McBRIDE, 2001). Esses dados estão de acordo com a imunolocalização encontrada, pois a região da bolsa paraflagelar é uma região de endocitose e juntamente com o Complexo de Golgi, existe uma intensa atividade de tráfego vesicular. Em *L. donovani* esta proteína está imunolocalizada em endossomos primários (MAROTTA *et al.*, 2006). Nos resultados de microarranjo e *western blot* encontramos um aumento na forma intermediária aderido por 24 horas. ARARIPE *et al* (2005) sugerem, por RT-PCR, que esse gene encontra-se aumentado na forma epimastigota, em relação a tripomastigotas. No entanto, esse estudo se baseou na realização de RT-PCR semi-quantitativa a partir de RNA celular total, o que impede a comparação com nossos dados, basados em mRNA polissomal.

O gene TcExp02B07 codifica uma proteína que possui um domínio roadblock/lc7, característicos de proteínas do complexo multimérico da dineína. A dineína pode ser classificada como citoplasmática ou axonemal e o domínio em questão faz parte das dineínas citoplasmáticas, as quais podem ser divididas em dois grupos, de acordo com suas funções:

- **Dineína citoplasmática do tipo 1:** envolvida em diversas funções como organização dos centríolos e migração nuclear durante a mitose, além do transporte de vesículas endocíticas e lisossomais,

- **Dineína citoplasmática do tipo 2:** envolvida em transporte intraparaflagelar, necessário para o funcionamento e montagem de cílios e flagelos (COLE, 2003; PFISTER *et al.*, 2006).

O domínio LC7 está classificado como dineína citoplasmática do tipo 2. O resultado de imunolocalização mostra áreas concentradas em forma de grânulos ao longo do flagelo e também um padrão difuso pelo citoplasma em menor intensidade. A imunomicroscopia eletrônica poderá dar uma melhor localização desta proteína que juntamente com ensaios de super-expressão poderão resultar em uma melhor caracterização no *T. cruzi*.

A proteína referente ao gene TcExp02B10 possui três domínios ligadores de cálcio, sendo característico da sub-unidade B da calcineurina. Esta proteína participa da transdução de sinal modulada por cálcio e calmodulina, sendo composta por duas sub-unidades de tamanhos distintos, A (unidade catalítica, ~60 kDa) e B (unidade regulatória, ~19 kDa) (MORIYA *et al* 1995; MONDRAGON *et al.*, 1997;). Na forma epimastigota de *T. cruzi*, a imunolocalização desta proteína aparece em duas regiões de forma consistente: na extremidade flagelo com maior intensidade de fluorescência e grânulos com uma intensidade menor em torno do núcleo.

Devido às funções diversas que dependem deste tipo de sinalização por cálcio, experimentos funcionais como inibidores de serina-treonina fosfatase juntamente com a imunolocalização por microscopia eletrônica poderão resultar em melhor comprovação dos resultados encontrados neste trabalho.

O gene TcExp02B12 apresentou um padrão de imunolocalização por microscopia óptica muito semelhante ao encontrado para o gene TcExp02B01, localizando-se entre a bolsa paraflagelar e o cinetoplasto. Este gene apresenta um domínio denominado ArfGap (*ADP-Ribosilation Factor/GTP-ase Activating Protein*) com função de hidrolisar GTP ligado à outra molécula denominada ARF, a qual está envolvida na regulação do tráfego de membranas, remodelamento de actina e formação de vesículas revestidas (RANDAZZO & HIRSCH; 2003). Da mesma forma que o gene TcExp02B01, a localização encontrada é compatível com a posição da via endocítica e/ou tráfego vesicular em *T. cruzi*.

A imunolocalização da proteína referente ao gene TcExp02C01 apresentou um padrão de imunolocalização semelhante ao encontrado para os genes TcExp02B01 e TcExp02B12. Essa proteína possui função relacionada ao sistema de fusão de membranas de vesículas sendo classificada como sendo uma syntaxina.

Estas três proteínas ligadas ao sistema endocítico e tráfego vesicular (TcExp02B01, TcExp02B12 e TcExp02C01) serão alvo de estudos posteriores, mais aprofundados, de localização utilizando microscopia de transmissão, para verificar sua localização de forma mais precisa. Esse estudo será muito importante pois muito pouco é conhecido sobre o sistema endocítico em *T. cruzi*, não havendo nenhuma publicação sobre syntaxina em *T. cruzi*.

A proteína referente ao gene TcExp02C04 é denominada de centrina sendo uma proteína ligadora de cálcio presente no centróssomo, presente em todos os eucariotos e podendo assumir diversas funções (ZAMORA & MARSHALL, 2005). Essa proteína foi localizada no cinetoplasto da forma epimastigota de *T. cruzi*, porém este dado será investigado mais minuciosamente com microscopia de transmissão. Em *T. brucei* a TbCen1 está localizada no corpo basal e a estrutura bi-lobada próximo ao Complexo de Golgi. Devido à baixa conservação de sequência entre os tripanosomatídeos, as funções podem variar entre os gêneros e espécies (SELVAPANDIYAN *et al.*, 2007).

É interessante notar que os genes TcExp02B12 e TcExp02C01 fazem parte do mesmo grupo de co-expressão no transcriptoma e que apresentaram o mesmo padrão de co-localização na imunofluorescência. Devido ao pequeno número de grupos para os quais foi possível imunolocalizar mais do que uma proteína, esse resultado demonstra que é possível identificar associações entre genes co-regulados, sendo que a ausência de outros exemplos é principalmente devida à dificuldade de se obter resultados completos para a maioria dos genes estudados. No entanto, o número de condições biológicas distintas avaliadas e utilizadas para a análise de co-expressão ainda é baixo, reforçando a necessidade, e importância, de que a base de dados transcriptômicos seja ampliada.

7 Conclusões

A partir do exposto nesse trabalho, pode-se concluir que:

- A seleção de genes candidatos para caracterização utilizando a técnica de microarranjo de DNA é uma abordagem produtiva, priorizando candidatos mais informativos para a análise de sua expressão, função e regulação, a partir de um conjunto mais amplo, se não o total, de proteínas expressas no organismo;
- Na era da genômica funcional, onde começamos a tentar compreender melhor a complexidade dos sistemas, torna-se necessário abordagens em larga escala visando elucidar os processos celulares, sendo que para isso devemos utilizar plataformas apropriadas. Nesse sentido, o sistema Gateway[®] mostrou-se eficaz na proposta de clonagem e expressão, podendo ser otimizada utilizando vetores desenhados especificamente. Sendo assim, pode-se afirmar que a metodologia escolhida representa uma escolha acertada para a construção do ORFeoma de *T. cruzi*;
- O resultado das análises comparando *western blot* com microarranjo de DNA apresentou um alto grau de correlação, comprovando a confiabilidade do método de análise utilizado no projeto de genômica funcional do IBMP e a relevância das modificações do transcriptoma na regulação da expressão gênica;
- A etapa de amplificação e clonagem dos genes foi obtida com sucesso em sua totalidade. Duas etapas apresentaram um menor índice de sucesso, expressão protéica e obtenção dos soros policlonais, as quais usualmente são gargalos em projetos desse tipo;
- O estudo dos genes selecionados proporcionou um incremento nas informações dos mesmos, com relação a aspectos bioinformáticos de anotação e de imunolocalização.
- Os indícios encontrados para o gene TcExp02A01 como sendo codificador de uma proteína ligadora de RNA e de estar localizado em corpúsculos citoplasmáticos possivelmente relacionados ao armazenamento de mRNA abrem perspectivas para um melhor aprofundamento do conhecimento sobre os mecanismos de regulação de expressão gênica pós-transcricional em *T. cruzi*, especificamente sobre esse gene.
- Dentre os objetivos práticos definidos para esse trabalho, obtivemos o seguinte:
 - Da clonagem de todos os genes na plataforma Gateway[®], não conseguimos somente um, devido a um erro na sequência do primer *forward*;

- Da expressão da metade do conjunto total de proteínas e de pelo menos sete proteínas hipotéticas, foram obtidas 19 proteínas expressas (67,8% do total) e 12 proteínas hipotéticas;
- Da imunolocalização de pelo menos cinco proteínas hipotéticas, obtivemos somente três. No entanto, há mais duas candidatas óbvias, que apresentaram sinal no *western* mas cuja imunofluorescência não ficou boa;
- Da obtenção de *western blot* quantitativo para pelo menos 10 proteínas, conseguimos obter onze anti-soros e, agregando os anti-soros de outros trabalhos, conseguimos a quantificação de 23 proteínas, todas realizadas nos mesmos extratos de *T. cruzi*, constituindo-se em uma base de dados essencial para a comprovação do impacto das modulações do transcriptoma nos níveis protéicos.
- Da análise funcional detalhada de pelo menos uma proteína hipotética, foi alcançando com a caracterização de TcExp02A01 como sendo uma proteína de ligação ao RNA, associada a corpúsculos *p-bodies-like*. A identificação dos seus alvos será feita no futuro próximo, complementando a sua caracterização funcional. Além disso, para esse objetivo foi identificado diversas proteínas associadas ao transporte vesicular, que em conjunto perfazem uma outra anotação funcional e que, após a caracterização de sua localização mais precisa, pelo uso de microscopia eletrônica de transmissão, completará a sua caracterização funcional mínima.

8 Perspectivas

Uma das características principais desse trabalho foi a necessidade de abranger um número relativamente grande de genes a serem caracterizados, com certeza acima da capacidade de serem analisados em um prazo temporal de dois anos, relativo à uma dissertação de mestrado. Ciente dessa limitação e não sendo possível, a priori, definir quais seriam os genes mais fáceis de serem caracterizados, optou-se por incluir diversos níveis de redundância e a definição de objetivos mínimos práticos, os quais, conforme mencionado acima, foram quase que totalmente cumpridos.

De qualquer forma, dentro dessa abordagem unidirecional, diversos elementos foram deixados durante o caminho e uma grande parte das perspectivas é relacionada aos mesmos:

- Expressão das proteínas que falharam nessa etapa, utilizando agora diversos protocolos, modificando especialmente a cepa bacteriana utilizada, o tempo e temperatura de indução e a concentração do indutor;
- Expressão de proteínas solúveis nos casos para os quais seria interessante fazer uma caracterização bioquímica de sua função;
- Melhorar a quantidade e a qualidade da informação obtida com os ensaios de *western blot*, ao se utilizar um conjunto mais amplo de amostras extraídas durante a metaciclogênese (epimastigota em fase de crescimento logarítmica, epimastigota em fase estacionária, epimastigotas estressados nutricionalmente, epimastigotas aderidos e em diferenciação por 3, 6, 12, 18, 24 horas aderidos e tripomastigota metacíclico), permitindo a identificação de padrões de modulação temporal mais precisos, e ao se quantificar o sinal da reação sob diversas perspectivas, aumentando a confiabilidade da mensuração.
- Realizar imunolocalização por microscopia óptica em um conjunto mais amplo de pontos da metaciclogênese (supracitados), em especial tripomastigotas metacíclicos, para as proteínas analisadas por imunofluorescência de epimastigotas no presente trabalho.
- Processar os anti-soros que apresentaram um padrão correto de identificação da proteína recombinante, mas que não reconheceram o extrato de *T. cruzi*.
- Realizar imunolocalização por microscopia eletrônica das proteínas que apresentaram um padrão adequado de imunolocalização por microscopia óptica.

- Com relação ao gene TcExp02A01, realizar ensaio de EMSA, imunoprecipitação, ribonômica, e co-localização com marcadores específicos.

9 Referências bibliográficas

- ABUIN, G.; FREITAS-JUNIOR, L.H.; COLLI, W.; ALVES, M.J. SCHENKMAN, S. Expression of trans-sialidase and 85-kDa glycoprotein genes in *Trypanosoma cruzi* is differentially regulated at the post-transcriptional level by labile protein factors. **Journal of Biological Chemistry**, v. 274, n. 19, p. 13041-13047, 1999.
- AEBERSOLD, R.; MANN, M. Mass spectrometry-based proteomics. **Nature**, v.422.p. 198-207, 2003.
- AGUERO, F.; VERDUN, R.E.; FRASCH, A.C.; SANCHEZ, D.O. A random sequencing approach for the analysis of the *Trypanosoma cruzi* genome: general structure, large gene and repetitive DNA families, and gene discovery. **Genome Research**, v. 10, n. 12, p. 1996-2005, 2000.
- AKOPYANTS, N.S.; MATLIB, R.S.; BUKANOVA, E.N.; SMEDS, M.R.; BROWNSTEIN, B.H.; STORMO, G.D.; BEVERLEY, S.M. Expression profiling using random genomic DNA microarrays identifies differentially expressed genes associated with three major developmental stages of the protozoan parasite *Leishmania major*. **Molecular and Biochemical Parasitology**, v. 136, n. 1, p. 71-86, 2004.
- ALMEIDA, R.; GILMARTIN, B.J.; McCANN, S.H.; NORRISH, A.; IVENS, A.C.; LAWSON, D.; LEVICK, M.P.; SMITH, D.F. et al. Expression profiling of the *Leishmania* life cycle: cDNA arrays identify developmentally regulated genes present but not annotated in the genome. **Molecular and Biochemical Parasitology**, v. 136, n. 1, p. 71-86, 2004
- ANDERSEN, J.S.; LYON, C.E.; FOX, A.H.; LEUNG, A.K.L.; LAM, Y.W.; STEEN, H.; MANN, M.; LAMOND, A.I. Directed proteomic analysis of the human nucleolus. **Current biology**, 12: 1-12, 2002.
- ARARIPE, J.R.; RAMOS, F.P.; CUNHA E SILVA, N.L.; URMÉNYI, T.P.; SILVA, R.; LEITE FONTES, C.F.; DA SILVEIRA, J.F.; RONDINELLI, E. Characterization of a RAB5 homologue in *Trypanosoma cruzi*. **Biochemical and Biophysical Research Communications**, v. 329, n. 2, p. 638-645, 2005.
- ARAVIND, L.; IYER, L.M.; ANANTHARAMAN, V. The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. **Genome Biology**, v. 4, n. 10, p. R64, 2003.
- ASLUND, L.; CARLSSON, L.; HENRIKSSON, J.; RYDAKER, M.; TORO, G.C.; GALANTI, N.; PETERSSON, U. A gene family encoding heterogeneous histone H1 protein in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 65, p. 317-330, 1994.
- ATWOOD, J.A. 3rd; WEATHERLY, D.B.; MINNING, T.A.; BUNDY, B.; CAVOLA, C; OPPERDOES, F.R.; ORLANDO, R.; TARLETON, R.L. The *Trypanosoma cruzi* proteome. **Science**, v. 309, n. 5733, p. 473-476, 2005.
- ÁVILA, A.R.; YAMADA-OGATA, S.F.; DA SILVA MONTEIRO, V.; KRIEGER, M.A.; NAKAMURA, C.V.; DE SOUZA, W.; GODENBERG, S. Cloning and characterization of the metacyclogenin gene, which is

- specifically expressed during *Trypanosoma cruzi* metacyclogenesis. **Molecular and Biochemical Parasitology**, v. 117, p. 169-177, 2001.
- ÁVILA, A.R.; DALLAGIOVANNA, B.; YAMADA-OGATTA, S.F.; MONTEIRO-GÓES, V.; FRAGOSO, S.P.; KRIEGER, M.A.; GOLDENBERG, S. Stage-specific gene expression during *Trypanosoma cruzi* metacyclogenesis. **Genetics and Molecular Research**, v. 2, n. 1, p. 159-168, 2003.
- AYMERICH, S.; GOLDENBERG, S.; The karyotype of *Trypanosoma cruzi* Dm 28c: comparison with other *T. cruzi* strains and Trypanosomatids. **Experimental Parasitology**, v. 69, p. 107-115, 1989.
- BANG, H.; PECHT, A.; RADDATZ, G.; SCIOR, T.; SOLBACH, W.; BRUNE, K.; PAHL, A. Prolyl isomerases in a minimal cell. Catalysis of protein folding by trigger factor from *Mycoplasma genitalium*. **European Journal of Biochemistry**, v. 267, n. 11, p. 3270-3280, 2000.
- BAPTISTA, C.S.; VENCIO, R.Z.; ABDALA, S.; VALADARES, M.P.; MARTINS, C.; BRAGANÇA PEREIRA, C.A.; ZINGALES, B. DNA microarrays for comparative genomics and analysis of gene expression in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 138, n. 2, p. 183-194, 2004.
- BAPTISTA, C.S.; VÊNCIO, R.Z.; ABDALA, S.; CARRANZA, J.C.; WESTENBERGER, S.J.; SILVA, M.N.; PEREIRA, C.A.; GALVÃO, L.M.; GONTIJO, E.D.; CHIARI, E.; STURM, N.R.; ZINGALES, B. Differential transcription profiles in *Trypanosoma cruzi* associated with clinical forms of Chagas disease: Maxicircle NADH dehydrogenase subunit 7 gene truncation in asymptomatic patient isolates. **Molecular and Biochemical Parasitology**, v. 150, n. 2, p. 236-248, 2006.
- BASSLER, M.; LINHARTOVÁ, I.; HALADA, P.; NOVOTNÁ, J.; BEZOUSKOVÁ, S.; OSICKA, R.; WEISER, J.; VOHRADSKÝ, J.; SEBO, P. The iron-regulated transcriptome and proteome of *Neisseria meningitidis* serogroup C. **Proteomics**, v. 6, n. 23, p. 6194-6206, 2006.
- BELL, S.D.; BOTTING, C.H.; WARDLEWORTH, B.N.; JACKSON, S.P.; WHITE, M.F. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. **Science**, v. 296, n. 5565, p. 148-151, 2002.
- BENTE, M.; HARDER, S.; WIESGIGL, M.; HEUKESHOVEN, J.; GELHAUS, C.; KRAUSE, E.; CLOS, J.; BRUCHHAUS, I.; Developmentally induced changes of the proteome in the protozoan parasite *Leishmania donovani*. **Proteomics**, 3: 1811-1829, 2003.
- BERBEROF, M.; VANHAMME, L.; TEBABI, P.; PAYS, A.; JEFFERIES, D.; WELBURN, S.; PAYS, E. The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. **Embo Journal**, v. 14, p. 2925-2934, 1995.
- BERRIMAN, M.; GHEDIN, E.; HERTZ-FOWLER, C. *et al.* The genome of the African trypanosome *Trypanosoma brucei*. **Science**, v. 309, n. 5733, p. 416-422, 2005.
- BLATTNER, J.; CLAYTON, C.E. The 3'-untranslated regions from the *Trypanosoma brucei* phosphoglycerate kinase-encoding genes mediate developmental regulation. **Gene**, v. 162, n. 1, p. 153-156, 1995.

- BLAXTER, M.; IVENS, A. Reports from the Cutting Edge of Parasitic Genome Analysis. **Parasitology Today**, v. 15, n. 11, p. 430-431, 1999.
- BLEMINGS, K.P.; CRENSHAW, T.D.; SWICK, R.W.; BENEVENGA, N.J. Lysine- α -ketoglutarate reductase and saccharopine dehydrogenase are located only in the mitochondrial matrix in rat liver. **The Journal of Nutrition**, v. 124, n. 8, p. 1215–1221, 1994.
- BONALDO, M.C., SOUTO-PADRON, T., DE SOUZA, W., AND GOLDENBERG, S. Cell-substrate adhesion during *Trypanosoma cruzi* differentiation. **Journal of Cellular Biology**, v. 106, p. 1349-1358, 1988.
- BONALDO, M.C.; D'ESCOFFIER, L.N.; SALLES, J.M.; GOLDENBERG, S. Characterization and expression of proteases during *Trypanosoma cruzi* metacyclogenesis. **Experimental Parasitology**, v. 73, n. 1, p. 44-51, 1991.
- BOUCHER, N.; WU, Y.; DUMAS, C.; DUBE, M.; SERENO, D.; BRETON, M.; PAPADOPOULOU, B. A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. **Journal of Biological Chemistry**, v. 277, n. 22, p. 19511-19520, 2002.
- BOUVIER-NAVÉ, P.; HUSSELSTEIN, T.; BENVENISTE, P. Two families of sterol methyltransferases are involved in the first and the second methylation steps of plant sterol biosynthesis. **European Journal of Biochemistry**, v. 256, n. 1, p. 88-96, 1998.
- BREMS, S.; GUILBRIDE, D.L.; GUNDLESODJIR-PLANCK, D.; BUSOLD, C.; LUU, V.D.; SCHANNE, M.; HOHEISEL, J.; CLAYTON, C. The transcriptomes of *Trypanosoma brucei* Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. **Molecular and Biochemical Parasitology**, v. 139, n. 2, p. 163-172, 2005.
- BRENER, Z.; ANDRADE, Z.A.; BARRAL-NETTO, M. *Trypanosoma cruzi e a Doença de Chagas* 2ed. Editora Guanabara-Koogan 2000.
- BRISSE, S.; BARNABE, C.; BANULS, A.L.; SIDIBE, I.; NOEL, S.; TIBAYRENC, M. A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. **Molecular and Biochemical Parasitology**, v. 92, n. 2, p. 253-263, 1998.
- BRISSE, S.; HENRIKSSON, J.; BARNABE, C.; DOUZERY, E.J.; BERKVENS, D.; SERRANO, M.; DE CARVALHO, M.R.; BUCK, G.A.; DUJARDIN, J.C.; TIBAYRENC, M. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. **Infection, Genetics and Evolution**, v. 2, n. 3, p. 173-183, 2003.
- BURLEIGH, B.A.; WOOLSEY, A.M. Cell signalling and *Trypanosoma cruzi* invasion. **Cell Microbiology**. V 4, n 11, p. 701-711, 2002.
- CAMARGO, E.P. Growth and differentiation in *Trypanosoma cruzi*. **Revista do Instituto de Medicina Tropical**, v. 6, p. 93-100, 1964.

- CANO, M.I.; GRUBER, A.; VAZQUEZ, M. *et al.* Molecular karyotype of clone CL Brener chosen for the *Trypanosoma cruzi* genome project. **Molecular and Biochemical Parasitology**, v. 71, n. 2, p. 273-278, 1995.
- CASSOLA, A.; DE GAUDENZI, J.G.; FRASCH, A.C. Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. **Molecular Microbiology**, v. 65, n. 3, p. 655-670, 2007.
- CASTELLANI, D.; RIBEIRO, L.V.; FERNANDES, J.F. Differentiation of *Trypanosoma cruzi* in culture. **Journal of Protozoology**, v. 14, n. 3, p. 447-451, 1967.
- CAUSTON, H.C.; REN, B.; KOH, S.S.; HARBINSON, C.T.; KANIN, E.; JENNINGS, E.G.; LEE, T.I.; TRUE, H.L.; LANDER, E.S.; YOUNG, A.R. Remodeling of yeast genome expression in response to environmental changes. **Molecular Biology of the Cell**, v. 12, p. 323-337, 2001.
- CHAGAS, C. Nova tripanosomíase humana. Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n.gen., n.sp., agente etiológico de nova entidade morbida do homem. **Memórias do Instituto Oswaldo Cruz**, v. 1, p. 159-218, 1909.
- CHUNG, S.H.; GILLESPIE, R.D.; SWINDLE, J. Analyzing expression of the calmodulin and ubiquitin-fusion genes of *Trypanosoma cruzi* using simultaneous, independent dual gene replacements. **Molecular and Biochemical Parasitology**, v. 63, p. 57-107, 1994.
- CLAYTON, C.E. Life without transcriptional control? From fly to man and back again. **EMBO Journal**, v. 21, p. 1881-1888, 2002.
- COHEN, A.M.; RUMPEL, K.; COOMBS, G.H.; WASTILING, J.M. Characterisation of global protein expression by two-dimensional electrophoresis and mass spectrometry: proteomics of *Toxoplasma gondii*. **International Journal for Parasitology**, v. 32, p. 39-51, 2002.
- COLE, D.G. The intraflagellar transport machinery of *Chlamydomonas reinhardtii*. **Traffic**, v. 4, n. 7, p. 435-442, 2003.
- CONTRERAS, V.T.; ARAUJO-JORGE, T.C.; BONALDO, M.C.; THOMAS, N.; BARBOSA, H.S.; MEIRELLES, M.N.L. & GOLDENBERG, S. Biological aspects of the Dm28c clone of *Trypanosoma cruzi* after metacyclogenesis in chemically defined media. **Memórias do Instituto Oswaldo Cruz**, v. 83, p. 123-133, 1988.
- CONTRERAS, V.T.; SALLES, J.M.; THOMAS, N.; MOREL, C.M. & GOLDENBERG, S. *In vitro* differentiation of *Trypanosoma cruzi* under chemically defined conditions. **Molecular and Biochemical Parasitology**, v. 16, p. 315-327, 1985.
- COOPER, R.A.; CARUCCI, D.J.; Proteomic approaches to studying drug targets and resistance in *Plasmodium*. **Current Drug Targets of Infectious Disorders**, v. 4, n. 1, p. 41-51, 2004.
- CORNELISSEN, A.W.; BACKES, S.; EVERS, R.; GRONDAL, E.J.; JESS, W.; KOCK, J. Transcription analysis in *Trypanosoma brucei*. **Biochemistry Society Transactions**, v. 18, n. 5, p. 710-714, 1990.

- CORNELISSEN, A.W.; EVERS, R.; GRONDAL, E.J.; HAMMER, A.; JESS, W.; KOCK, J. Transcription and RNA polymerases in *Trypanosoma brucei*. **Acta Leiden**, v; 58, n. 2, p. 75-96, 1989.
- COUGHLIN, B.C.; TEIXEIRA, S.M.; KIRCHHOFF LV, DONELSON, J.E. Amastin mRNA abundance in *Trypanosoma cruzi* is controlled by a 3'-untranslated region position-dependent cis-element and an untranslated region-binding protein. **Journal of Biological Chemistry**, v. 251, n. 16, p. 12051-12060, 2000.
- COURA, J.R.; DE CASTRO, S.L. A critical review on Chagas disease chemotherapy. **Memórias do Instituto Oswaldo Cruz**, v. 97, n. 1, p. 3-24, 2002.
- DALLAGIOVANNA, B.; PLAZANET-MENUT, C.; OGATTA, S.F.Y.; ÁVILA, A.R; KRIEGER, M.A.; GOLDENBERG, S. *Trypanosoma cruzi*: a gene family encoding chitinbinding-like proteins is post-transcriptionally regulated during metacyclogenesis. **Experimental Parasitology**, v. 99, p. 7-16, 2001.
- DAGANZO, S.M.; ERZBERGER, J.P.; LAM, W.M.; SKORDALAKES, E.; ZHANG, R.; FRANCO, A.A.; BRILL, S.J.; ADAMS, P.D.; BERGER, J.M.; KAUFMAN, P.D. Structure and function of the conserved core of histone deposition protein Asf1. **Current Biology**, v. 13, n. 24, p. 2148-2158, 2003.
- DAS, A.; BELLOFATTO, V. RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor. **Proceedings of the National Academy of Sciences, USA**, v. 100, n. 1, p. 80-85, 2003.
- DE ANDRADE, A.F.; ESTEVES, M.J.; ANGLUSTER, J.; GONZALES-PERDOMO, M.; GOLDENBERG, S. Changes in cell-surface carbohydrates of *Trypanosoma cruzi* during metacyclogenesis under chemically defined conditions. **Journal of General Microbiology**, v. 137, n. 12, p. 2845-2849, 1991.
- DE SOUZA, W. Cell Biology of *Trypanosoma cruzi*. **International Reviews in Cytology**, v. 86, p. 197-283, 1984.
- DEGRAVE, W.M.; MELVILLE, S.; IVENS, A.; AASLETT, M. Parasite genomes initiatives. **International Journal for Parasitology**, v. 31, p. 532-536, 2001.
- DHAND, R. Functional genomics. **Nature**, v. 405, p. 819, 2000.
- DIEHL, S.; DIEHL, F.; EL-SAYED, N.M.; CLAYTON, C.E.; HOHEISEL, J.D. Analysis of stage-specific gene expression in the bloodstream and the procyclic form of *Trypanosoma brucei* using a genomic DNA microarray. **Molecular and Biochemical Parasitology**, v. 123, n. 2, p. 115-123, 2002.
- DIETRICH, P.; SOARES, M.B.; AFFONSO, M.H.; FLOETER-WINTER, L.M. The *Trypanosoma cruzi* ribosomal RNA-encoding gene: analysis of promoter and upstream intergenic spacer sequences. **Gene**, v. 125, n. 1, p. 103-107, 1993.
- DIKENG, A.; FERREIRA, L.; D'ANGELO, M.; DOLEZAL, P.; LAMB, T.; MURTA, S.; TRIGGS, V.; ULBERT, S.; VILLARINO, A.; RENZI, S.; ULLU, E; TSCHUDI, C. Characterization of a candidate *Trypanosoma brucei* U1 small nuclear RNA gene. **Molecular and Biochemical Parasitology**, v. 133, n. 1, p. 109-115, 2001.

- DOERKS, T.; COPLEY, R.R.; SCHULTZ, J.; PONTING, C.P.; BORK, P. Systematic identification of novel protein domain families associated with nuclear functions. **Genome Research**, v. 12, n. 1, p. 47-56, 2002.
- D'ORSO, I.; DE GAUDENZI, J.G.; FRASCH, A.C. RNA-binding proteins and mRNA turnover in trypanosomes. **Trends in Parasitology**, v. 19, n. 4, p.151-155, 2003.
- DRAGON, E.A.; SIAS, S.R.; KATO, E.A.; GABE, J.D. The genome of *Trypanosoma cruzi* contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. **Molecular and Cellular Biology**, v. 7, p. 275-280, 1987.
- DRUMMELSMITH, J.; BROCHU, V.; GIRARD, I.; MESSIER, N.; OUELLETTE, M. Proteome mapping of the Protozoan parasite *Leishmania* and application to the study of drug targets and resistance mechanisms. **Molecular and Cellular Proteomics**, v. 2, n. 3, p. 146-155, 2003.
- DUHAGON, M.A.; DALLAGIOVANNA, B.; CIGANDA, M.; RUYECHAN, W.; WILLIAMS, N. GARAT, B. A novel type of single-stranded nucleic acid binding protein recognizing a highly frequent motif in the intergenic regions of *Trypanosoma cruzi*. **Biochemical and Biophysical Research Communications**, v. 309, n. 1, p. 183-188, 2003.
- DUSANIC, D.G. *In vitro* production of metacyclic tripomastigotes. **Journal of Parasitology**, v. 66, n. 6, p. 1046-1049, 1980.
- EDMAN, P. A method for the determination of the amino acid sequence of peptides. **Archives of Biochemistry and Biophysics**, v. 22, p. 475-483, 1949.
- EISENBERG, D.; MARCOTTE, E.M.; XENARIOS, I.; YEATES, T.O. Protein function in the post-genomic era. **Nature**, v. 405, p. 823-826, 2000.
- ELIAS, M.C.; MARQUES-PORTO, R.; FREYMULLER, E.; SCHENKMAN, S. Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. **Molecular and Biochemical Parasitology**, v. 112, p. 79-90, 2001.
- ELIAS, M.C.Q.B.; FARIA, M.; MORTARA, R.A.; MOTTA, M.C.M.;de SOUZA, W.; THIRY, M.; SCHENKMAN, S. Chromosome localization changes in the *Trypanosoma cruzi* nucleus. **Eukaryotic Cell**, v. 1, n. 6, p. 944-953, 2002.
- EL-SAYED, N.M.; MYLER, P.J.; BARTHOLOMEU, D.C. *et al.* The genome sequence of *Trypanosoma cruzi* etiologic agent of Chagas disease. **Science**, v. 309, n. 5733, p. 409-415, 2005.
- EL-SAYED, N.M.; MYLER, P.J.; BLANDIN, G. *et al.* Comparative genomics of trypanosomatid parasitic protozoa. **Science**, v. 309, n. 5733, p. 404-409, 2005.
- ENGLUND, P.T.; GUILBRIDE, L.; HWA, K.Y.; JOHNSON, C.E.; LI, C.; ROCCO L.J.; TORRI, A.F. Kinetoplast DNA: structure and replication. **Molecular Biology of Parasitic Protozoa**, p. 75-87, 1996.
- ESTEVEZ, M.G.; GONZALES-PERDOMO, M.; ALVIANO, C.S.; ANGLUSTER, J.; GOLDENBERG, S. Changes in fatty acid composition associated with differentiation of *Trypanosoma cruzi*. **FEMS Microbiologic Letters**, v. 50, n. 1-2, p. 31-34, 1989.

- EVERS, R.; HAMMER, A.; KOCK, J.; JESS, W.; BORST, P.; MEMET, S.; CORNELISSEN, A.W. *Trypanosoma brucei* contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. **Cell**, v. 56, n. 4, p. 585-597, 1989.
- FAKHRY, Y.E.; OUELLETTE, M.; PAPADOPOULOU, B. A proteomic to identify developmentally regulated proteins in *Leishmania infantum*. **Proteomics**, v. 2, p. 1007-1017, 2002.
- FASSHAUER, D.; SUTTON, R.B.; BRUNGER, A.T.; JAHN, R. Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. **Proceedings of the National Academy of Sciences, USA**, v. 95, p. 15781-15786, 1998.
- FIGUEIREDO, R.C.; ROSA, D.S.; SOARES, M.J. Differentiation of *Trypanosoma cruzi* epimastigotes: metacyclogenesis and adhesion to substrate are triggered by nutritional stress. **Journal of Parasitology**, v. 86, n. 6, p. 1213-1218, 2000.
- FLORENS, L.; WASHBURN, M.P.; RAINE, J.D.; ANTHONY, R.M.; GRAINGER, M.; HAYNES, J.D.; MOCH, J.K.; MUSTER, N. SACCI, J.B.; TABB, D.L.; WITNEY, A.A.; WOLTERS, D.W.U.Y.; GARDNER, M.J.; HOLDER, A.A.; SINDEN, R.E.; YATES, J.R.; CARUCC, I.D.J. A proteomic view of the Plasmodium falciparum life cycle. **Nature**, v. 419, n. 6906, p. 520-526, 2002.
- FRAGOSO, S.P.; PLAZANET-MENUT, C.; CARREIRA, M.A.; MOTTA, M.C.; DALLAGIOVANNA, B.; KRIEGER, M.A.; GOLDENBERG, S. Cloning and characterization of a gene encoding a putative protein associated with U3 small nuclear ribonucleoprotein in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 126, p. 113-117, 2003.
- FRAIDENRAICH, D.; PENA, C.; ISOLA, E.L.; LAMMEL, E.M.; COSO, O.; ANEL, A.D.; PONGOR, S.; BARALLE, F.; TORRES, H.N.; FLAWIA, M.M. Stimulation of *Trypanosoma cruzi* adenylyl cyclase by an alpha D-globin fragment from Triatoma hindgut: effect on differentiation of epimastigote to trypomastigote forms. **Proceedings of the National Academy of Sciences, USA**, v. 90, n. 21, p. 10140-10144, 1993.
- FUTCHER, B.; LATTE, G.I.; MONARDO, P.; McLAUGHIN, C.S.; GARRELS, J.I. A sampling of the yeast proteome. **Molecular and Cellular Biology**, v. 19, n. 11, p. 7357-7368, 1999.
- GARCIA, E.S.; AZAMBUJA, P.; Development and interactions of Trypanosoma cruzi within the insect vector. **Parasitol Today**. Sep;7(9):240-4,1991.
- GARCIA, E.S.; GONZALEZ, M.S.; DE AZAMBUJA, P.; BARALLE, F.E.; FRAIDENRAICH, D.; TORRES, H.N.; FLAWIA, M.M. Induction of Trypanosoma cruzi metacyclogenesis in the gut of the hematophagous insect vector, Rhodnius prolixus, by hemoglobin and peptides carrying alpha D-globin sequences. **Experimental Parasitology**, v. 81, n. 3, p. 255-261, 1995.
- GAUNT, M.W.; YEO, M.; FRAME, I.A.; TOTHARD, J.R.; CARRASCO, H.J.; TAYLOR, M.C.; MENA, S.S.; VEAZEY, P.; MILES, G.A.; ACOSTA, N.; ROJAS DE ARIAS, A.; MILES, M.A. Mechanism of genetic exchange in American trypanosomes. **Nature**, v.421, p. 936-939, 2003.

- GILINGER, G.; BELLOFATTO, V. Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. **Nucleic Acids Research**, v. 29, n. 7, p. 1556-1564, 2001.
- GOLDENBERG, S.; SALLES, J.M.; CONTRERAS, V.T.; LIMA FRANCO, M.P.A.; KATZIN, A.M.; COLLI, W. & MOREL, C.M. Characterization of messenger RNA from epimastigotes and metacyclic trypomastigotes of *Trypanosoma cruzi*. **FEBS**, v. 180, n.2, p. 265-270, 1985.
- GOLDENBERG, S. *Trypanosoma cruzi* as a model system for studying cell differentiation. **Parasites: Molecular Biology, Drug and Vaccine Design**, p. 1-6, 1990.
- GOLDENBERG, S. Gene Expression during development of *Trypanosoma cruzi*, C. C. Wang, ed. (Washington, DC: American Association for the Advancement of Science), 1991.
- GÓMEZ, E.B.; MEDINA, G.; BALLESTA, J.P.; LEVIN, M.J.; TELLEZ-INON, M.T. Acidic ribosomal P proteins are phosphorylated in *Trypanosoma cruzi*. **International Journal of Parasitology**, v.31, n. 10, p. 1032-1039, 2001.
- GONZALES, A.; LERNER, T.J.; HUECAS, M.; SOSA-PINEDA, B.; NOGUEIRA, N.; LISARD, P.M. Apparent generation of a segment RNAm from two separate tandem families in *Trypanosoma cruzi*. **Nucleic Acids Research**, v. 13, p. 5789-5804, 1985.
- GONZALES-PERDOMO, M.; ROMERO, P.; GOLDENBERG, S. Cyclic AMP and adenylate cyclase activators stimulate *Trypanosoma cruzi* differentiation. **Experimental Parasitology**, v.66, n. 2, p. 205-212, 1988.
- GRAVES, P.R.; HAYSTEAD, T.A.J. Molecular biologist's guide to proteomics. **Microbiology and Molecular Biology Reviews**, v. 66, n. 1, p. 39-63, 2002.
- GRIFFIN, T.J.; GYGI, S.P.; IDEKER, T.; RIST, B.; ENG, J.; HOOD, L.; AEBERSOLD, R. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. **Molecular and Cellular Proteomics**, v. 1, n. 4, p. 323-333, 2002.
- GROS, L.; CASTILLO-ACOSTA, V.M.; JIMÉNEZ JIMÉNEZ, C.; SEALEY-CARDONA, M.; VARGAS, S.; MANUEL ESTÉVEZ, A.; YARDLEY, V.; RATTRAY, L.; CROFT, S.L.; RUIZ-PEREZ, L.M.; URBINA, J.A.; GILBERT, I.H.; GONZÁLEZ-PACANOWSKA, D. New azasterols against *Trypanosoma brucei*: role of 24-sterol methyltransferase in inhibitor action. **Antimicrobial Agents Chemotherapy**, v. 50, n. 8, p. 2595-601, 2006.
- GUIMOND, C.; TRUDEL, N.; BROCHU, C.; MARQUIS, N.; EL FADILI, A.; PEYTAVI, R *et al.* Modulation of gene expression in *Leishmania* drug resistant mutants as determined by targeted DNA microarrays. **Nucleic Acids Research**, v. 31, n. 20, p. 5886-5896, 2003.
- GUO, R.; XUE, H.; HUANG, L. Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. **Molecular Microbiology**, v. 50, n. 5, p. 1605-1615, 2003.
- GYGI, S.P.; ROCHON, Y.; FRANZA, B.R. AEBERSOLD, R. Correlation between protein and mRNA abundance in yeast. **Molecular and Cellular Biology**, v. 19, p. 1720-1730, 1999.

- HAJDUK, S.L.; SABATINI, R.S. RNA editing: post-transcriptional restructuring of genetic information. **Molecular Biology of Parasitic Protozoa**, p. 135-158, 1996.
- HARTLEY, J.L.; TEMPLE, G.F.; BRASCH, M.A. DNA cloning using in vitro site-specific recombination. **Genome Research**, v. 10, n. 11, p. 1788-1795, 2000.
- HEHL, A.; VASSELLA, E.; BRAUN, R.; RODITI, I. A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in *Trypanosoma brucei*. **Proceedings National Academy Sciences USA**, v. 91, n. 1, p. 370-374, 1994.
- HENRIKSON, J.; PORCEL, B.; RYDAKER, M.; RUIZ, A.; CAZZULO, J.J.; FRASCH, A.C.C.; PETERSSON, U.; Chromosome specific markers reveal conserved, linkage groups in spite of extensive chromosomal size variation in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 37, p. 64-73, 1995.
- HOARE, C.A.; WALLACE, F.G.; Developmental stages of trypanosomatid flagellate: a new terminology. **Nature**, v. 244, p. 69-70, 1966.
- HOLETZ, F.B.; CORREA, A.; ÁVILA, A.R.; NAKAMURA, C.V.; KRIEGER, M.A.; GOLDENBERG, S. Evidence of P-body-like structures in *Trypanosoma cruzi*. **Biochemical and Biophysical Research Communications**, v. 356, n. 4, p. 1062-1067, 2007.
- HOLZER, T.R.; McMASTER, W.R.; FORNEY, J.D. Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in *Leishmania mexicana*. **Molecular and Biochemical Parasitology**, v. 146, n. 2, p. 198-218, 2006.
- HOMSY, J.J.; GRANGER, B.; KRASSNER, S.M. Some factors inducing formation of metacyclic stages of *Trypanosoma cruzi*. **Journal of Protozoology**, v. 36, n. 2, p. 150-153, 1989.
- HUH, W.K.; FALVO, J.V.; GERKE, L.C.; CARROLL, A.S.; HOWSON, R.W.; WEISSMAN, J.S.; O'SHEA, E.K. Global analysis of protein localization in budding yeast. **Nature**, v. 425, n. 6959, p. 686-91, 2003.
- IDEKER, T.; THORSSON, V.; RANISH, J.A.; CHRISTMAS, R.; BUHLER, J.; ENG, J.K.; BUMGAMER, R.; GOODLETT, D.R.; AEBERSOLD, R.; HOOD, L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. **Science**, v. 292, n. 5518, p. 929-934, 2001.
- ISOLA, E.L.; LAMMEL, E.M.; GONZALEZ CAPPA, S.M. *Trypanosoma cruzi*: differentiation to metacyclic trypomastigotes in the presence of ADP-ribosyltransferase inhibitors. **Experimental Parasitology**, v.64, n.3, 424-429, 1987.
- IVENS, A.C.; PEACOCK, C.S.; WORTHEY, E.A. *et al.* The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v.309, n. 5733, p.436-442, 2005.
- JÄRVINEN, A.K.; HAUTANIEMI, S.; EDGREN, H.; AUVINEN, P.; SAARELA, J.; KALLIONIEMI, O.P.; MONNI, O. Are data from different gene expression microarray platforms comparable? **Genomics**, v. 83, n. 6, p. 1164-1168, 2004.

- JESS, W.; HAMMER, A.; CORNELISSEN, A.W. Complete sequence of the gene encoding the largest subunit of RNA polymerase I of *Trypanosoma brucei*. **FEBS Letters**, v.249, n.1, p. 123-128, 1989.
- KLEFFMANN, T.; SCHMIDT, J.; SCHAUB, G.A. Attachment of *Trypanosoma cruzi* epimastigotes to hydrophobic substrates and use of this property to separate stages and promote metacyclogenesis. **Journal of Eukaryotic Microbiology**, v.45, n.5, p.548-555, 1998.
- KOCK, J.; EVERS, R.; CORNELISSEN, A.W. Structure and sequence of the gene for the largest subunit of trypanosomal RNA polymerase III. **Nucleic Acids Research**, v.16, n.18, p.8753-8772, 1988.
- KOCK, J.; CORNELIASSEN, A.W. Characterization of the RNA polymerases of *Crithidia fasciculata*. **Molecular Microbiology**, v.5, n.4, p.835-842, 1991.
- KOLLIEN, A.H.; SCHAUB, G.A. The development of *Trypanosoma cruzi* in triatominae. **Parasitology Today**, v. 16, p. 381-387, 2000.
- KRASSNER, S.M.; GRANGER, B.; LEE, P.; GUERRA, C.; LE, T.; LUC, K.O. Action of exogenous potassium and calcium ions on *in vitro* metacyclogenesis of *Trypanosoma cruzi*. **Journal of Protozoology**, v.38, n.6, p.602-608, 1991.
- KRAUSE, M.; HIRSH, D. A trans-spliced leader sequence on actin mRNA in *C. elegans*. **Cell**, v.49, n.6, p.753-761, 1987.
- KUMAR, A.; AGARWAL, S.; HEYMAN, J.A.; MATSON, S.; HEIDTMAN, M.; PICCIRILLO, S.; UMANSKY, L.; DRAWID, A.; JANSEN, R.; LIU, Y.; CHEUNG, K.H.; MILLER, P.; GERSTEIN, M.; ROEDER, G.S.; SNYDER, M. Subcellular localization of the yeast proteome. **Genes and Development**, v. 16, n. 6, p. 707-719, 2002.
- LAIRD, P.W. Trans *splicing* in trypanosomes-archaism or adaptation? **Trends in Genetics**, v.5, n.7, p.204-208, 1989.
- LANDFEAR, S.M. Trypanosomatid transcription factors: waiting for Godot. **Proceedings of the National Academy of Sciences, USA**, v.100, n.1, p.7-9, 2003.
- LANDY, A. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. **Annual Review of Biochemistry**, v. 58, p. 913-949, 1989.
- LANZETTI, L.; PALAMIDESSI, A.; ARECES, L.; SCITA, G.; DI FIORE, P.P. Rab5 is a signalling GTPase involved in actin remodelling by receptor tyrosine kinases. **Nature**, v. 429, n. 6989, p. 309-314, 2004.
- LEBER, R.; LANDL, K.; ZINSER, E.; AHORN, H.; SPOK, A.; KOHLWEIN, S.D.; TURNOWSKY, F.; DAUM, G. Dual localization of squalene epoxidase, Erg1p, in yeast reflects a relationship between the endoplasmic reticulum and lipid particles. **Molecular and Cellular Biology**, v. 2, p. 375-86, 1998.
- LEBOWITZ, J.H.; SMITH, H.Q.; RUSCHE, L.; BEVERLEY, S.M. Coupling of poly(A) site selection and trans-*splicing* in *Leishmania*. **Genes and Development**, v.7, n.6, p. 996-1007, 1993.

- LEIFSO, K.; COHEN-FREUE, G.; DOGRA, N.; MURRAY, A.; McMASTER, W.R. Genomic and proteomic expression analysis of *Leishmania* promastigote and amastigote life stages: the *Leishmania* genome is constitutively expressed. **Molecular and Biochemical Parasitology**, v. 152, n. 1, p. 35-46, 2007.
- LEVINE, N. D.; CORLISS, J. O.; COX, F.E.G.; DEROUX, G.; GRAIN, J.; HONIGBERG, B.M.; LEEDADE, G. F.; LOEBLICH, A.R.; LOM, F.; LYNN, D.; MERINFELD, E. G.; PAGE, F.C.; POLJANSKI, G.; SPRAGUE, V.; VAVRA, J. & WALLACE, F.C. A newly revised classification of the Protozoa. **Journal of Protozoology**, v.27, p.37-58, 1980.
- LIU, S.; MILNE, G.T.; KUREMSKY, J.G.; FINK, G.R.; LEPLA, S.H. Identification of the proteins required for biosynthesis of diphthamide, the target of bacterial ADP-ribosylating toxins on translation elongation factor 2. **Molecular and Cellular Biology**, v. 24, n. 21, p. 9487-97, 2004.
- LOCKHART, D.J.; WINZELER, E.A. Genomics, gene expression and DNA arrays. **Nature**, v. 405, p. 827-836, 2000.
- LUKES, J.; GUILBRID, L.; VOTÝPKA, J.; ZIKOVÁ, A.; BENNE, R.; ENGLUND, P.T. Kinetoplast DNA Network: evolution of an improbable structure. **Eukaryotic Cell**, v. 4, n. 1, p. 495-502, 2002.
- MACHADO, C.A.; AYALA, F.J. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. **Proceedings of the National Academy of Sciences, USA**, v.98, p.7396-7401, 2001.
- MACINA, R.A.; SANCHEZ, D.O.; AFFRANCHINO, J.C.; ENGEL, J.; FRASCH, A.C.C. Polymorphism within minicircle sequence classes in the kDNA of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 16, p. 61-74, 1985.
- MAINGON, R.; GERK, R.; RODRIGUEZ, M.; URBINA, J.; HOENICKA, J.; NEGRI, S.; AGUIRRE, T.; NEHLIN, J.; KNAPP, T.; CRAMPTON, J. The Tubulin genes of *Trypanosoma cruzi*. **European Journal of Chemistry**, v. 264, p. 285-291, 1988.
- MAIR, G.; SHI, H.; LI, H.; DJIKENG, A.; AVILES, H.O.; BISHOP, J.R.; FALCONE, F.H.; GAVRILESCU, C.; MONTGOMERY, J.L.; SANTORI, M.I.; STERN, L.S.; WANG, Z.; ULLU, E.; TSCHUDI, C. A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. **RNA**, v. 6, n. 2, 163-169, 2000.
- MAKAROVA, K.S.; ARAVIND, L.; KOONIN, E.V. SWIM, a novel Zn-chelating domain present in bacteria, archaea and eukaryotes. **Trends in Biochemical Sciences**, v. 27, p. 384-386, 2002.
- MARIN-NETO, J.A.; ALMEIDA FILHO, O.C.; PAZIN FILHO, A.; MACIEL, B.C. Indeterminate form of Chagas' disease. Proposal of new diagnostic criteria and perspectives for early treatment of cardiomyopathy. **Arquivos Brasileiros de Cardiologia**, v. 79, n. 6, p. 623-627, 2002.
- MAROTTA, D.E.; GERALD, N.; DWYER, D.M. Rab5b localization to early endosomes in the protozoan human pathogen *Leishmania donovani*. **Molecular and Cellular Biochemistry**, v. 292, n. 1-2, p. 107-117, 2006.

- MARSH, V.L.; PEAK-CHEW, S.Y.; BELL, S.D. Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. **Journal of Biological Chemistry**, v. 280, n. 22, p. 21122-21128, 2005.
- MASLOW, D.A.; SIMPSOM, L. Evolution of parasitism in Kinetoplastid protozoa. **Parasitology Today**, v. 11, n. 1, p. 30-32, 1995.
- MATTHEWS, K.R.; TSCHUDI, C.; ULLU, E. A common pyrimidine-rich motif governs trans-*splicing* and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. **Genes and Development**, v.8, n.4, p.491-501, 1994.
- McBRIDE, A.E.; COOK, J.T.; STEMMLER, E.A.; RUTLEDGE, K.L.; McGRATH, K.A.; RUBENS, J.A. Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions. **Journal of Biological Chemistry**, v. 280, n. 35, p. 30888-30898, 2005.
- McCARTHY-BURKE, C.; TAYLOR, Z.A.; BUCK, G.A. Characterization of the spliced leader genes and transcripts in *Trypanosoma cruzi*. **Gene**, v.82, n.1, 177-189, 1989.
- McDONAGH, P.D.; MYLER, P.J.; STUART, K. The unusual gene organization of *Leishmania major* chromosome 1 may reflect novel transcription processes. **Nucleic Acids Research**, v.28, n.14, p.2800-2803, 2000.
- McNICOLL, F.; DRUMMELSMITH, J.; MULLER, M.; MADORE, E.; BOILARD, N.; OUELETTE, M.; PAPADOPOULOU, B. A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*. **Proteomics**, v. 6, n. 12, p. 3567-3581, 2006.
- MIJALSKI, T.; HARDER, A.; HALDER, T.; KERSTEN, M.; HORSCH, M.; STROM, T.M.; LIEBSCHER, H.V.; LOTTSPEICH, F.; DE ANGELIS, M.H.; BECKERS, J. Identification of coexpressed gene clusters in a comparative analysis of transcriptome and proteome in mouse tissues. **Proceedings of the National Academy of Sciences, USA**, v. 102, n. 24, p. 8621-8626, 2005.
- MINNING, T.A.; BUA, J.; GARCIA, G.A.; McGRAW, R.A.; TARLETON, R.L. Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 131, n. 1, p. 55-64, 2003.
- MOGGS, J.G.; ORPHANIDES, G. Genomic analysis of stress response genes. **Toxicology Letters**, v. 140-141, p. 149-53, 2003.
- MONDRAGON, A.; GRIFFITH, E.C.; SUN, L.; XIONG, F.; ARMSTRONG, C.; LIU, J.O. Overexpression and purification of human calcineurin alpha from *Escherichia coli* and assessment of catalytic functions of residues surrounding the binuclear metal center. **Biochemistry**, v. 36, n. 16, p. 4934-4942, 1997.
- MONNERAT, S.; MARTINEZ-CALVILLO, S.; WORTHEY, E.; MYLER, P.J.; STUART, K.D.; FASEL, N. Genomic organization and gene expression in a chromosomal region of *Leishmania major*. **Molecular and Biochemical Parasitology**, v. 134, n. 2, p. 233-243, 2004.

- MORIYA, M.; FUJINAGA, K.; YAZAWA, M.; KATAGIRI, C. Immunohistochemical localization of the calcium/calmodulin-dependent protein phosphatase, calcineurin, in the mouse testis: its unique accumulation in spermatid nuclei. **Cellular Tissue Research**, v. 281, n. 2, p. 273-281, 1995.
- MUNAKATA, T.; ADACHI, N.; YOKOYAMA, N.; KUZUHARA, T.; HORIKOSHI, M.A. A human homologue of yeast anti-silencing factor has histone chaperone activity. **Genes and Cells**, v. 5, n. 3, p. 221-233, 2000.
- NARDELLI, S.C.; AVILA, A.R.; FREUND, A.; MOTTA, M.C.; MANHÃES, L.; DE JESUS, T.C.; SCHENKMAN, S.; FRAGOSO, S.P.; KRIEGER, M.A.; GOLDENBERG, S.; DALLAGIOVANNA, B. Small-subunit rRNA processome proteins are translationally regulated during differentiation of *Trypanosoma cruzi*. **Eukaryotic cell**, v. 6, n. 2, p. 337-345, 2007.
- NILSEN, T.W. Trans-splicing in protozoa and helminths. **Infectious Agents and Diseases**, v.1, n.4, p.212-218, 1992.
- NILSSON, D.; ANDERSSON, B. Strand asymmetry patterns in trypanosomatid parasites. **Experimental Parasitology**, v.109, n.3, p.143-149, 2005.
- NISSON, P.M.; SANNY, A.; KOK, Y.J.; HIANG, Y.T.; CHUAH, S.H.; SHING, T.K.; LEE, Y.Y.; WONG, K.T.; HU, W.S.; SIM, M.Y.; PHILP, R. Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. **Molecular Biotechnology**, v. 34, n. 2, p. 125-140, 2006.
- NOGUEIRA, N.; BIANCO, C.; COHN, Z. Studies on the selective lysis and purification of *Trypanosoma cruzi*. **Journal of Experimental Medicine**, v. 142, n. 1, 224-229, 1975.
- NORRIS, K.A.; BRADT, B.; COOPER, N.R. SO, M. Characterization of a *Trypanosoma cruzi* C3 binding protein with functional and genetic similarities to the human complement regulatory protein, decay-accelerating factor. **J Immunol**. Oct 1;147(7):2240-71991.
- NOZAKI, T.; CROSS, G.A. Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v.75, n.1, p.55-67, 1995.
- NUGENT, P.G.; KARSANI, S.A.; WAIT, R.; TEMPERO, J.; SMITH, D.F. Proteomic analysis of *Leishmania mexicana* differentiation. **Molecular and Biochemical Parasitology**, v. 136, n. 1, p. 51-62, 2004.
- NUNES, L.R.; CARVALHO, M.R.; SHAKARIAN, A.M.; BUCK, G.A. The transcription promoter of the spliced leader gene from *Trypanosoma cruzi*. **Gene**, v.188, n.2, p.157-168, 1997.
- O'ROURKE, N.A.; MEYER, T.; CHANDY, G. Protein localization studies in the age of 'Omics'. **Current Opinion in Chemical Biology**, v. 9, p. 82-87, 2005.
- OCHATT, C.M.; ULLOA, R.M.; TORRES, H.N.; TELLEZ-INON, M.T. Characterization of the catalytic subunit of *Trypanosoma cruzi* cyclic AMP-dependent protein kinase. **Molecular and Biochemical Parasitology**, v.57, n.1, p.73-81, 1993.
- OVERATH, P.; HAAG, J.; LISCHKE, A.; O'HUIGIN, C. The surface structure of trypanosomes in relation to their molecular phylogeny. **International Journal of Parasitology**, v.31, n.5-6, p.468-471, 2001.

- PABA, J.; RICART, C.A.O.; FONTES, W.; SANTANA, J.M.; TEIXEIRA, A.R.L.; MARCHESE, J.; WILLIAMSON, B.; HUNT, T. KARGER, B.L.; SOUZA, M.V. Proteomic Analysis of *Trypanosoma cruzi* Developmental Stages Using Isotope-Coded Affinity Tag Reagents. **Journal of Proteome Research**, v. 3, n. 3, p. 517-524, 2003.
- PABA, J.; SANTANA, J.M.; TEIXEIRA, A.R.L.; FONTES, W.; SOUZA, M.V.; RICART, C.A.O. Proteomic analysis of the human pathogen *Trypanosoma cruzi*. **Proteomics**, v. 4, n. 4, p. 1052-1059, 2004.
- PALFI, Z.; LANE, W.S.; BINDEREIF, A. Biochemical and functional characterization of the cis-spliceosomal U1 small nuclear RNP from *Trypanosoma brucei*. **Molecular and Biochemical Parasitology**, v.121, n.2, p.233-243, 2002.
- PALFI, Z.; SCHIMANSKI, B.; GUNZL, A.; LUCKE, S.; BINDEREIF, A. U1 small nuclear RNP from *Trypanosoma brucei*: a minimal U1 snRNA with unusual protein components. **Nucleic Acids Research**, v.33, n.8, p.2493-2503, 2005.
- PANDEY, A.; MANN, M. Proteomics to study genes e genomes. **Nature**, v. 405, p. 837-846, 2000.
- PARODI-TALICE, A.; DÚRAN, R.; ARRAMBIDE, N.; PRIETO, V.; PINERO, M.D.; PRITSCH, O.; CAYOTA, A.; CERVENANSKY, C.; ROBELLO, C. Proteome analysis of the causative agent of chagas disease: *Trypanosoma cruzi*. **International Journal for Parasitology**, v. 34, n. 8, p. 881-886, 2004.
- PATTERSON, S.D.; AEBERSOLD R.H. Proteomics: the first decade and beyond. **Nature Genetics**, Supplement, v. 33, p. 311-323, 2003.
- PAWLAK, M.R.; SCHERER, C.A.; CHEN, J.; ROSHON, M.J.; RULEY, H.E. Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. **Molecular and Cellular Biology**, v. 20, n. 13, p. 4859-4869, 2000.
- PETERSEN, D.; CHANDRAMOULI, G.V.; GEOGHEGAN, J.; HILBURN, J.; PAARLBERG, J.; KIM, C.H.; MUNROE, D.; GANGI, L.; HAN, J.; PURI, R.; STAUDT, L.; WEINSTEIN, J.; BARRETT, J.C.; GREEN, J.; KAWASAKI, E.S. Three microarray platforms: an analysis of their concordance in profiling gene expression. **BMC Genomics**, v. 6, n. 1, p. 63, 2005.
- PFISTER, K.K.; SHAH, P.R.; HUMMERICH, H.; RUSS, A.; COTTON, J.; ANNUAR, A.A.; KING, S.M.; FISHER, E.M. Genetic Analysis of the Cytoplasmic Dynein Subunit Families. **PLoS Genetics**, v. 2, n. 1, p. 1, 2006.
- PROBST, C.M. Descrição da Metaciclologênese de *Trypanosoma cruzi* pelo uso de microarranjos de DNA. Tese de Doutorado apresentada ao Instituto Oswaldo Cruz, 2005.
- PROUDFOOT, N.; FURGER, A.; DYE, M. Integrating mRNA processing with transcription. **Cell**, v.108, n.4, p.501-512, 2002.
- RAJKOVIC, A.; DAVIS, R.E.; SIMONSEN, J.N.; ROTTMAN, F.M. A spliced leader is present on a subset of mRNAs from the human parasite *Schistosoma mansoni*. **Proceedings of the National Academy of Sciences, USA**, v. 87, n. 22, p. 8879-8883, 1990.

- RANDAZZO, P.A.; HIRSCH, D.S. Arf GAPs: multifunctional proteins that regulate membrane traffic and actin remodelling. **Cell Signaling**, v. 16, n. 4, p. 401-413, 2004.
- RECINOS, R.F.; KIRCHHOFF, L.V.; DONELSON, J.E. Cell cycle expression of histone genes in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 113, n. 2, p. 215-222, 2001.
- REQUENA, J.M.; LOPEZ, M.C.; JIMENEZ-RUIZ, A.; DE LA TORRE, J.C.; ALONSO, C. A head-to-tail tandem organization of hsp70 genes in *Trypanosoma cruzi*. **Nucleic Acids Research**, v. 16, p. 1393-1406, 1988.
- RODRIGUES-POUSADA, C.A.; NEVITT, T.; MENEZES, R.; AZEVEDO, D.; PEREIRA, J.; AMARAL, C. Yeast activator and stress response: an overview. **FEBS Letters**, v. 567, n. 1, p. 80-85, 2004.
- RUAL J.F.; HILL, D.E.; VIDAL, M. ORFeome projects: gateway between genomics and omics. **Current Opinion in Chemical Biology**, v. 8, p. 20-25, 2004.
- RUIZ, F.A.; RODRIGUES, C.O.; DoCAMPO, R. Rapid changes in polyphosphate content within acidocalcisomes in response to cell growth, differentiation, and environmental stress in *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 276, n. 28, p. 26114-26121, 2001.
- SALMON, D.; MONTERO-LOMELI, M.; GOLDENBERG, S. A DnaJ-like protein homologous to the yeast co-chaperone Sis1 (TcJ6p) is involved in initiation of translation in *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 276, n. 47, p. 43970-43979, 2001.
- SANDMAN, K.; REEVE, J.N. Archaeal chromatin proteins: different structures but common function? **Current Opinion in Microbiology**, v. 8, n. 6, p. 656-661, 2005.
- SAXENA, A.; WORTHEY, E.A.; YAN, S.; LELAND, A.; STUART, K.D.; MYLER, P.J. Evaluation of differential gene expression in *Leishmania major* Friedlin procyclics and metacyclics using DNA microarray analysis. **Molecular and Biochemical Parasitology**, v. 129, n. 1, p. 103-114, 2003.
- SAXENA, A.; LAHAV, T.; HOLLAND, N.; AGGARWAL, G.; ANUPAMA, A.; HUANG, Y.; VOLPIN, H.; MYLER, P.J.; ZILBERSTEIN, D. Analysis of the *Leishmania donovani* transcriptome reveals an ordered progression of transient and permanent changes in gene expression during differentiation. **Molecular and Biochemistry Parasitology**, v. 152, n. 1, p. 53-65, 2007.
- SBICEGO, S.; ALFONZO, J.D.; ESTEVEZ, A.M.; RUBIO, M.A.; KANG, X.; TURCK, C.W.; PERIS, M.; SIMPSON, L. RBP38, a novel RNA-binding protein from trypanosomatid mitochondria, modulates RNA stability. **Eukaryotic Cell**, v. 2, n. 3, p. 560-568, 2003.
- SCHAUB, G.A. Metacyclogenesis of *Trypanosoma cruzi* in the vector *Triatoma infestans*. **Memórias do Instituto Oswaldo Cruz**, v. 83, p. 563-570, 1988.
- SCHENA, M.; SHALON, D.; DAVIS, R.W.; BROWN, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science**, 270(5235):467-70, 1995.

- SCHERL, A.; FRANÇOIS, P.; CHARBONNIER, Y.; DESHUSSES, J.M.; KOESSLER, T.; HUYGHE, A. *et al.* Exploring glycopeptide-resistance in *Staphylococcus aureus*: a combined proteomics and transcriptomics approach for the identification of resistance-related markers. **BMC Genomics**, v. 7, p. 296, 2006.
- SCHMIDT, J.; KLEFFMAN, T.; KOLLIEN, A.; SCHAUB, G.A. Further studies on hydrophobic attachment of *T. cruzi* epimastigotas. . **Memórias do Instituto Oswaldo Cruz**, v.92 (suppl):78, 1997.
- SCHWEDER, T.; HECKER, M. Monitoring of stress responses. **Advances in Biochemistry Engineering and Biotechnology**, v. 89, p. 47-71, 2004.
- SELVAPANDIYAN, A.; KUMAR, P.; MORRIS, J.C.; SALISBURY, J.L.; WANG, C.C.; NAKHASI, H.L. Centrin1 is required for organelle segregation and cytokinesis in *Trypanosoma brucei*. **Molecular Biology of the Cell**, 2007
- SICKMANN, A.; REINDERS, J.; WAGNER, Y.; JOPPICH, C.; ZAHEDI, R.; MEYER, H.E.; SCHONFISCH, B.; PERSCHIL, I.; CHACINSKA, A.; GUIARD, B.; REHLING, P.; PFANNER, N.; MEISINGER, C. The proteome of *Saccharomyces cerevisiae* mitochondria. **Proceedings of the National Academy of Sciences, USA**. V. 100, n. 23, p. 13207-13212, 2003.
- SIMPSON, L. The mitochondrial genome of Kinetoplastid protozoa: genomic organization, transcription, replication and evolution. **Annual Review of Microbiology**, v. 41, p. 363-382, 1987.
- SMITH, J.L.; LEVIN, J.R.; AGABIAN, N. Molecular characterization of the *Trypanosoma brucei* RNA polymerase I and III largest subunit genes. **Journal of Biological Chemistry**, v.264, n.30, p.18091-18099, 1989.
- SOARES, M.J. The reservosome of *Trypanosoma cruzi* epimastigotes: an organelle of the endocytic pathway with a role on metacyclogenesis. **Memórias do Instituto Oswaldo Cruz**, v.94, p.139-141, 1999.
- SOARES, M.J.; SOUTO-PADRON, T.; BONALDO, M.C.; GOLDENBERG, S.; DE SOUZA, W. A stereological study of the differentiation process in *Trypanosoma cruzi*. **Parasitology Research**, v.75, n.7, p.522-527, 1989.
- SOUZA, D.W. Basic Cell Biology of *Trypanosoma cruzi*. **Current Pharmaceutical Design**, v. 8, p. 269-285, 2002.
- SRIVIDYA, G.; DUNCAN, R.; SHARMA, P.; RAJU, B.V.; NAKHASI, H.L.; SALOTRA, P. Transcriptome analysis during the process of in vitro differentiation of *Leishmania donovani* using genomic microarrays. **Parasitology**, p. 1-13, 2007.
- STEVENS, J.R.; NOYES, H.A.; SCHOFIELD, C.J.; GIBSON, W. The molecular evolution of Trypanosomatidae. **Advances in Parasitology**, v.48, p.1-56, 2001.
- STUART, K. RNA editing: an overview, status report, and personal perspective. **Molecular Approaches to Parasitology**, p. 243-244, 1995.
- SUTTON, R.E.; BOOTHROYD, J.C. Evidence for trans *splicing* in trypanosomes. **Cell**, v.47, n.4, p.527-535, 1986.

- SWINDLE, J.; AJIOKA, J.; EISEN, H.; SAWAL, B.; JACQUEMOT, C.; BROWDER, Z.; BUCK, G. The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. **EMBO Journal**, v. 7, p. 1121-1127, 1988.
- TDR, 2005 Seventeenth Programme Report Progress 2003-2004; www.who.int/tdr/publications/publications/pdf/pr17/chagas.pdf, acesso em 28/11/2007.
- TEIXEIRA, S.M.; RUSSELL, D.G.; KIRCHHOFF, L.V.; DONELSON, J.E. A differentially expressed gene family encoding "amastin," a surface protein of *Trypanosoma cruzi* amastigotes. **Journal Biological Chemistry**, v. 269, n. 32, p. 20509-20516, 1994.
- TEIXEIRA, S.M.; KIRCHHOFF, L.V.; DONELSON, J.E. Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 270, n. 38, p. 22586-22594, 1995.
- TEIXEIRA, S.M.; DAROCHA, W.D. Control of gene expression and genetic manipulation in the Trypanosomatidae. **Genetic and Molecular Research**, v. 2, n. 1, p. 148-58, 2003.
- TESSIER, L.H.; KELLER, M.; CHAN, R.L.; FOURNIER, R.; WEIL, J.H.; IMBAULT, P. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-splicing in *Euglena*. **EMBO Journal**, v. 10, n. 9, p. 2621-2625, 1991.
- THONGBOONKERD, V.; BARATI, M.T.; McLEISH, K.R.; PIERCE, W.M.; EPSTEIN, P.M.; KLEIN, J.B. Proteomics and diabetic nephropathy. **Contributions in Nephrology Basel Karger**, v. 141, p. 142-145, 2004.
- TOMÁS, A.M.; KELLY, J.M. Stage-regulated expression of cruzipain, the major cysteine protease of *Trypanosoma cruzi* is independent of the level of RNA. **Molecular and Biochemical Parasitology**, v. 76, p. 91-103, 1996.
- TONELLI, R.R.; SILBER, A.M.; ALMEIDA-DE-FARIA, M.; HIRATA, I.Y.; COLLI, W.; ALVES, M.J. L-proline is essential for the intracellular differentiation of *T. cruzi*. **Cell Microbiology**, v.6, n.8, p.733-741, 2004.
- TOWBIN, H.; STAEBELIN, T.; GORDON, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. **Proceedings of the National Academy of Sciences, USA**, v. 76, n. 9, p. 4350-4354, 1979.
- TSCHUDI, C. Transcription and RNA processing in *Trypanosoma brucei*. **Molecular Approaches to Parasitology**, p. 255-268, 1995.
- TYERS, M.; MANN, M. From genomics to proteomics. **Nature**, v. 422, p. 193-197, 2003.
- TYLER-CROSS, R.E.; SHORT, S.L.; FLOETER-WINTER, L.M.; BUCK, G.A. Transient expression mediated by the *Trypanosoma cruzi* rRNA promoter. **Molecular and Biochemical Parasitology**, v. 72, n. 1-2, p. 23-31, 1995.

- ULLOA, R.M.; MESRI, E.; ESTEVA, M.; TORRES, H.N.; TELLEZ-INON, M.T. Cyclic AMP-dependent protein kinase activity in *Trypanosoma cruzi*. **Biochemical Journal**, v. 255, n. 1, p. 319-326, 1988.
- ULLU, E.; TSCHUDI, C. Trans *splicing* in trypanosomes requires methylation of the 5' end of the spliced leader RNA. **Proceedings of the National Academy of Sciences, USA**, v. 88, n. 22, p. 10074-10078, 1991.
- ULLU, E.; TSCHUDI, C.; GUNZI, A. Trans-*splicing* in trypanosomatid protozoa. **Molecular Biology of Parasitic Protozoa**, p. 115-129, 1996.
- UNWIN, R.D.; PIERCE, A.; WATSON, R.B.; STERNBERG, D.W.; WHETTON, A.D. Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. **Molecular and Cellular Proteomics**, v. 4, n. 7, p. 924-935, 2005.
- URBINA, J.A. Intermediary metabolism of *Trypanosoma cruzi*. **Parasitology Today**, v.10, n.3, p.107-110, 1994.
- VANHAMME, L.; PAYS, E. Control of gene expression in trypanosomes. **Microbiology Reviews**, v.59, p.223-240, 1995.
- VASQUEZ, M.P.; SCHIJMAN, A.G.; LEVIN, M.J. A short interspersed repetitive element provides a new 3' acceptor site for trans-*splicing* in certain ribosomal P2 β protein genes of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 64, p. 327-336, 1994.
- VICKERMAN, K.; TETLEY, L. Recent ultrastructural studies on trypanosomes. **Ann Soc Belg Med Trop**, v. 57, p. 441-457, 1977.
- VIDAL, M. A biological atlas of functional maps. **Cell**, v. 104, p. 333-339, 2001.
- WANG, H.; KACHMAN, M.T.; SCHWARTZ, D.R.; CHO, K.R.; LUBMAN, D.M. Comprehensive proteome analysis of ovarian cancers using liquid phase separation, mass mapping and tandem mass spectrometry: A strategy for identification of candidate cancer biomarkers. **Proteomics**, v. 4, n. 8, p. 2476-2495, 2004.
- WARD, P.; EQUINET, L.; PACKER, J.; DOERIG, C. Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. **BMC Genomics**, v.5, n.1, p.79, 2004.
- WASHBURN, M.P.; KOLLER, A.; OSHIRO, G.; ULASZEK, R.R.; PLOUFFE, D.; DECIU, C.; WINZELER, E.; YATES J.R. 3rd. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. **Proceedings of the National Academy of Sciences, USA**, v. 100, n. 6, p. 3107-3112, 2003.
- WEIMBS, T.; LOW, S.H.; CHAPIN, S.J.; MOSTOV, K.E.; BUCHER, P.; HOFMANN, K.A. A conserved domain is present in different families of vesicular fusion proteins: a new superfamily. **Proceedings of the National Academy of Sciences, USA**, v. 94, n. 7, p. 3046-3051, 1997.
- WILKINS, M.R.; PASQUALI, C.; APPEL, R.D.; OU, K.; GOLAZ, O.; SANCHEZ, J.C.; YAN, J.X.; GOOLEY, A.A.; HUGHES, G.; HUMPHERY-SMITH, I.; WILLIAMS, K.L.; HOCHSTRASSER, D.F. From proteins to proteome: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. **Biotechnology**, v. 14, n. 1, p. 61-65, 1996.

- WORTHEY, E.A.; SCHNAUFER, A.; MIAN, I.S.; STUART, K.; SALAVATI, R. Comparative analysis of editosome proteins in trypanosomatids. **Nucleic Acids Research**, v.31, n.22, p.6392-6408, 2003.
- XU, H.; ALGUINDIGUE, S.S.; WEST, A.H.; COOK, P.F. A proposed proton shuttle mechanism for saccharopine dehydrogenase from *Saccharomyces cerevisiae*. **Biochemistry**, v. 46, n. 3, p. 871-882, 2007.
- YAN, W.; LEE, H.; YI, E.C.; REISS, D.; SHANNON, P.; KWIECISZEWSKI, B.K.; COITO, C.; LI, X.J.; KELLER, A.; ENG, J.; GALITSKI, T.; GOODLETT, D.R.; AEBERSOLD, R.; KATZE, M.G. System-based proteomic analysis of the interferon response in human liver cells. **Genome Biology**, v. 5, n. 8, p. R54, 2004.
- YANG, Y.H.; SPEED, T. Design issues for cDNA microarray experiments. **Nature Reviews**, v. 3, p. 579-588, 2002.
- ZAMORA, I.; MARSHALL, W.F. A mutation in the centriole-associated protein centrin causes genomic instability via increased chromosome loss in *Chlamydomonas reinhardtii*. **BMC Biology**, v. 31, n. 3, p. 15, 2005.
- ZELEDON, R, ALVARENGA, N. J, SCHOSINSKY, K. In "*ChagasDisease*. **Pan American Health Organization**. Publ. 347 pp. 59. 1997.
- ZERIAL, M.; McBRIDE, H. Rab proteins as membrane organizers. **Nature Reviews of Molecular and Cellular Biology**, v. 2, n. 2, p. 107-117, 2001.
- ZHOU, X.W.; BLACKMAN, M.J.; HOWELL, S.A.; CARRUTHERS, V.B. Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. **Molecular and Cellular Proteomics**, v. 3, n. 6, p. 565-576, 2004
- ZHU, H.; BILGIN, M.; SNYDER, M. Proteomics. **The annual Review of Biochemistry**, 72:783-812, 2003.
- ZHU, X; TANG, G; GALILI, G. Characterization of the two saccharopine dehydrogenase isozymes of lysine catabolism encoded by the single composite AtLKR/SDH locus of *Arabidopsis*. **Plant Physiology**, v. 124, n. 3, p. 1363-1372, 2000.
- ZINGALES, B.; PEREIRA, M.E.; ALMEIDA, K.A.; UMEZAWA, E.S.; NEHME, N.S.; OLIVEIRA, R.P.; MACEDO, A.; SOUTO, R.P. Biological parameters and molecular markers of clone CL Brener the reference organism of the *Trypanosoma cruzi* genome project. **Memórias do Instituto Oswaldo Cruz**, v.92, n.6, p.811-814, 1997.
- ZWIERZYNSKI, T.A.; BUCK, G.A. RNA-protein complexes mediate *in vitro* capping of the spliced-leader primary transcript and U-RNAs in *Trypanosoma cruzi*. **Proceedings of the National Academy of Sciences, USA**, v.88, n.13, p.5626-5630, 1991.